

[illegible]

BASE COUNT	208	3	202	234	146	2 others
ORIGIN						

Query Match	32.38;	Score 457.8;	DB 9;	Length 792;
Best Local Similarity	97.78;	Pred. No. 1.2e-129;		
Matches 505;	Conservative	0;	Mismatches 12;	Indels 0;
			Gaps	0;

[illegible]

```

"organism": "Homo sapiens",
"/db_xref="taxon:9606",
"/clone="IMAGE:3690319",
"/clone_lib="NIL_Mer_44",
"/tissue_type="endometrium, adenocarcinoma",
"/lab_host="H1h1b (Phage-Resistant)",
"/note="Organ: uterus; Vector: pOTH7; Site:1"

```

Query Match	28.38;	Score	436.6;	EB	10;	Length	828;
Best Local Similarity	93.08;	Pred	No. 2.80-112;				
Matches	468;	Conservative	0;	Mismatches	34;	Indels	1;
						Gaps	1

[illegible][illegible]

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100																																																																																																																																																																													
Q7	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805	806	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859	860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876	877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899	900	901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918	919	920	921	922	923	924	925	926	927	928	929	930	931	932	933	934	935	936	937	938	939	940	941	942	943	944	945	946	947	948	949	950	951	952	953	954	955	956	957	958	959	960	961	962	963	964	965	966	967	968	969	970	971	972	973	974	975	976	977	978	979	980	981	982	983	984	985	986	987	988	989	990	991	992	993	994	995	996	997	998	999	1000

Accession	Sequence	Position
Db	705 TGTCCCAAGAGACACGGACATTTCTCCACAGGAACCTGTTACTTCATGTCCTAACTCCACGGG	764
2Y	908 AACTGCGACACACCTCCGTCACCCGC	930
Db	765 AACTGGGACGACACCTCCGTCACCCGC	787

DEFINITION	AU140165, PLACE2 Homo sapiens cDNA clone IMAGE20000625, mRNA
DESCRIPTION	172 bp
FEATURES	1..172 bp mRNA
ACCESSION	AU140165
VERSION	AU140165
KEYWORDS	EST.
SOURCE	human
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 792)	Okada, T., Suzuki, Y., Saito, K., Ishii, S., Yamamoto, J., Saitohama, T., Nishikawa, T., Nakamura, Y., Sugano, S., Masuhara, Y. and Isegai, T.	HRI human cDNA project (Okada, T., Suzuki, Y., Saito, K., Ishii, S., Yamamoto, J., Saitohama, T., Nishikawa, T., Nakamura, Y., Sugano, S., Masuhara, Y., Isegai, T.)	Unpublished (2000)	
			Conf. Fact: Takao Isegai	Genomics Laboratory

15323 Iida, Kitarazu, Chiba 292 0012, Japan

FEATURES	Location/Qualifiers
source	1. .792

Query Match	Score	DB	Length
Best Local Similarity	76.78	412.4	792
Matches	422	Conservative	0
		Mismatches	16
		Gaps	0

[illegible][illegible][illegible]

LOCUS	AV645755	706 bp	mpna	linear	EST 15-JAN-2002
DEFINITION	AV645755	GLC Homo sapiens cDNA clone	GLCAED05 3'		mRNA sequence.
ACCESSION	AV645755				
VERSION	AV645755.1				

SOURCE	ORGANISM
human	<i>homo sapiens</i>
Eukaryota: Metazoa:	Chordata: Stenidra: Vertebrata: Fishes: Osteichthys:
Mammalia: Eutheria:	Primates: Catarrhini: Hominoidea: Homo:
1 (bases 1 to 706)	
Xu, X., Huang, J., Xu, Z., Qian, R., Zhu, Z., Yan, Q., Cai, T., Zhang, X., Xiao, H., Qiu, T., Jin, F., Huang, Q., Cheng, Z., Li, N., Du, J., Hu, W., Shen, K., Lu, G., Fu, G., Zheng, W., Xu, S., Gu, W., Huang, W., Zhao, X.,	

by comparing gene expression profiles of hepatocellular carcinoma

With those of corresponding noncancerous liver
Proc Natl Acad Sci U S A. 98 (26): 15083-15094 (2001)
JOURNAL: Acad. Sci U S A. 98 (26): 15083-15094 (2001)
MEDLINE: 21625106
COMMENT: Zengyuan Han
Chinese National Human Genome Project at Shanghai

Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801932
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES	Location/Qualifiers
Source	1. . 706

```

/note="Vector: pBluescript sk(-), Site_1 EcoRI; Site_2
XhoI"
BASE COUNT      168 a      176 g      162 t      2 others
ORIGIN

```

Query Match	25.0%	Score	385.4	DB	10	Length	706
Best Local Similarity	99.2%	Pred	No	9	7e-98		
Matches	387	Conservative	0	Mismatches	3	Indels	0
						Gaps	0

OY	825	CATTGAAACCCCTGTGGCGGCAGTGTGCCAAGSACTTGATTTCTTCGAAGSAACGTGTT	884
Dd	140	CACAAGAAGCCTGTGCCCGCACGTGCCAAGGACTGACATTTCTTCGAAGSAACGTGTT	199
OY	885	ACTCATCTATTCACTCCACAGGGAAGTGGCAAGCATCGGTCATCCGCTTGGCAAGAAATGTA	944

Db 200 A C T C A G T C T A A C T C C C A C A G C G A A C T G G C C A C G A C T T C C T C A C C G T C T C A C G A A G T G A 253

Oy 945 GGGCCAGCTGTGTAATTAACCTGCTGAGGAGAGAACTTCTTACAGCTGCAGACCT 100

Db 260 GGGCCAGCTCGTCGTAATZAAACGGCTGAGGAGATGGAACCTTCCTACAGCTGGAGACTT 319

QY 1005 CCAGGAGCTAACCCGCTCTCTCTCGAATGGAGACTTCAGACTTAAATATAGTAAAGGATATCTATAT 106

Db 320 CAGGAGTAAACCCGCTCTCTCTGGATGGGACCTTTCAGAGCTAAATCAGTAAGGCTAACTGAGG 477

[illegible]

1125 GGACGATAGCGGGTGAAGACTGAGGTAATTAATCTATCTGTTGAACTAAAT 1136

440 CCAACATATGCGGATTAACATCTCTGGTAATTATATGCTATCTGAAATGACATC 499

QY 1185 GATGTGACGTTGACCAATTACTGATCTGCA 1214

Db 500 GATGTGACCTGACAACTACTGATCTGCA 529

RESULT 7

Accession	Length	Type	EST
R98113	503 bp	mrna	linear
LOCUS			EST 11-SEP

DEFINITION
Y966A02.11 Boates fedai liver spleen INF5 Homo sapiens cDNA IMAGE:200714 5' similar to SF:A46274 A46274 HIV GP120-BINDING

ACCESSION VERSION	C-YPE LECTIN - 1, mRNA sequence.
R98113	
R98113.1	CT-983773

VERSION	190115.1	01.2007
KEYWORDS	EST.	
SOURCE	human.	

ORGANISM	Homo sapiens
Eukaryota; Metazoa, Chordata, Craniata, Vertebrata; Euteleostei	

REFERENCE
1 (bases 1 to 503)
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

AUTHORS Hillier, L., Clark, N., Dubuque, J., Elliston, K., Hawkins, M., Holt

100

TITLE
JOURNAL
COMMENT

The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Address: Washington University School of Medicine
Washington University School of Medicine

Washington University School of Medicine
4444 Forest Park Parkway, Box 8001, St. Louis, MO 63110
Tel.: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1845
High quality sequence stops: 370
Source: IMAGE Consortium, LINT

This clone is available royalty-free through LNL : clone@lnc.com
 IMAGE Consortium (image.jhu.edu) for further info
 Insert Length: 1845 Std Error: 0.00
 Seq primer: M13P1
 High quality sequence stop: 370.

FEATURES	LOCATION/Qualifiers
SOURCE	1. .503
	/jordan.jsm="Homo sapiens"

```

3970 /ab_xref="GDB:376976.4"
3971 /ab_xref="taxon:9606"
3972 /clone="IMAGE:200714"
3973 /clone_id="Scars fetal liver spleen INFs"
3974 /sex="male"
3975 /dev_stage="20 week post conception fetus"
3976 /lab_host="DH10b (ampicillin resistant)"
3977 /note="Organ: Liver and Spleen; Vector: pT7
with a modified polylinker; Site_1: Pac I;
Site_2: NotI; The first strand cDNA was primed with a Pac I
[5'-AAGTCGACGAATTAAATGAAGCCTTTTTTTTTTTT-

```

BASE COUNT	126 a	137 c	128 d	195 f	7 others
ORIGIN					

Query Match	24.8%	Score 392.4	DB 14	Length 50
Best Local Similarity	97.0%	Prod. No. 4.9e-97		
Matches 419	Conservative	0	Mismatches	10

B-27 CATTCTAACTGTTTCTATTAATTTTCAGGTAAAG
GTT

DB 61 CACACGACGCCCTGTGGCCGCACTCTCCCTAAAGCTCTGGACATTCCTCTCCAAAGGAAAG

```

007 885 ACICAGGCTTAACCTTCTATCTATAATTATGACCTCTCTACGACGCGGCGACACGAG
      |||||

```

Dd 121 AGTCAGTCTCAACTGCGACGGGGAAATGGTAAGAATTATAGTTGTATATGAAGC

945 GGGGGCCAGCTCGTCTAATCAAAACATGCTGAGGAGCAGAACCTCCCTACAGCCTGGT
 2Y
 181 GGGGGGACCTGCTCTCTAATCAAAACCTGCTGAGGAGCAGAACTTCTTACAGCTCCGTA

[illegible]

241 CAGAGTAAACGGCTTCTCTGGATGAGCTTTCAGAGCTAAATCAGGAGAGCCATC

1065 AATGGTTCATACGCGTCTCTGACCGGCTTCCATGGTACTGGACACGCGG

Db 301 AATGGGTGGACGGCTCACTCTGTCACTCACTTCCAGGGTACTGGACACAGTCA

1125 CCACCAIACCCGATGACACTCG-CCGATTTAGTGGCACTGGCT GGAAG

Db 361 CCACCAATAGCGGGAATGAGACTGTGTTTAAATTAGTGGCTAGTGGCTGGCAAC

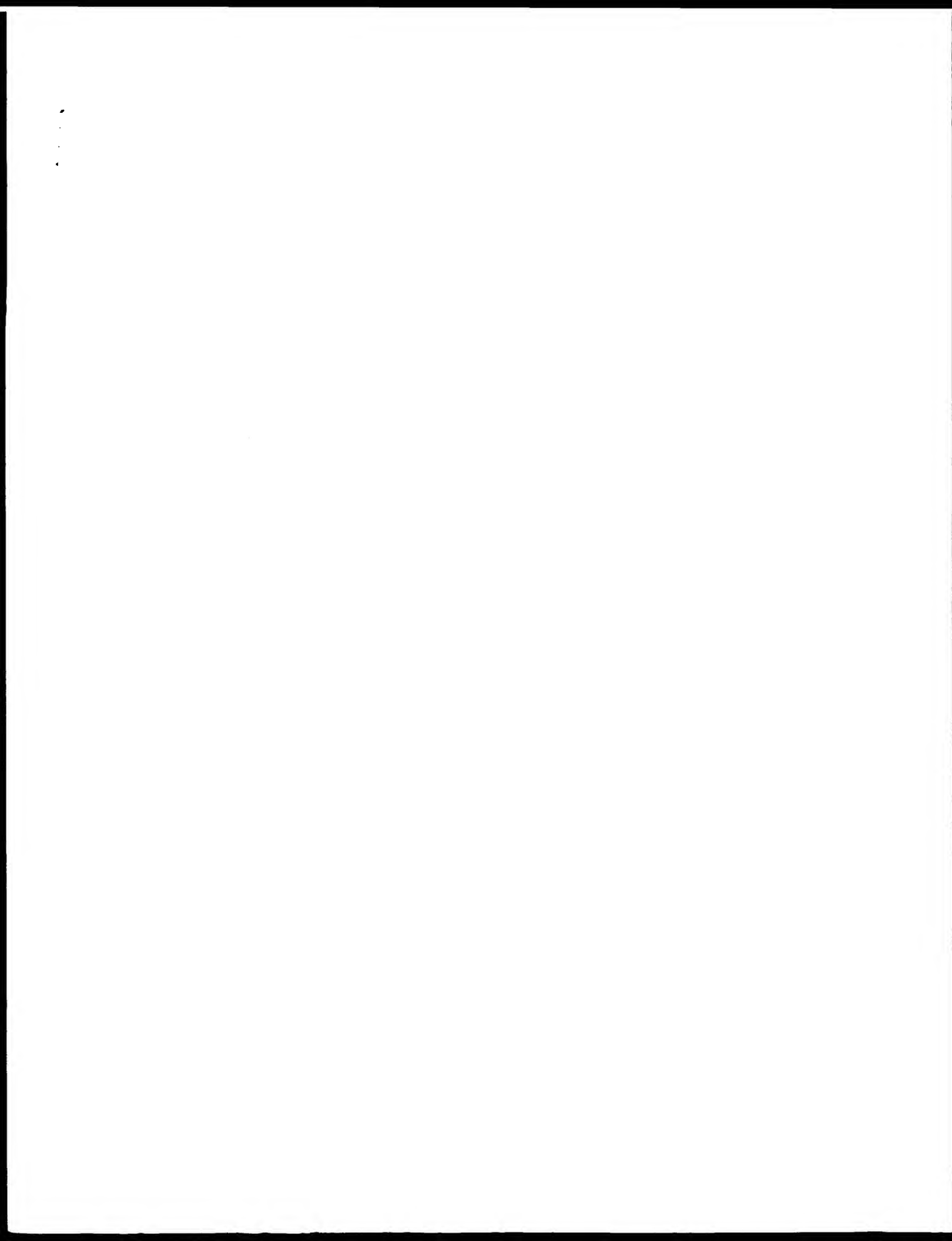
1183 TCGATGTCATGTTAATAATATCGATCTGAAAAAGCGCGAGCCCTGCTTCAGA
27 |||||

Db 421 TCGATGTGAC-TTNACATTACTGGATCTGTAAAAAGCCCGCAGCNIGTTTGGAA

BM386285/ 654 bp mRNA linear EST 17-2AN 2002
 LOCUS 01-R-CNI-cj-f-d-03-0-01.s1 01-R-CNI Rattus norvegicus cDNA clone
 DEFINITION 01-R-CNI-cj-f-d-03-0-01 3', mRNA sequence.
 ACCESSION BM386285
 VERSION BM386285 1 GI:18186334
 KEYWORDS EST.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Metazoa; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 654)
 Ronald, M. F., Lennon, G. and Soares, M. B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 97044477
 CONTACT: Soares, M.B.
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.iowa.edu
 The sequence contained an oligo-dT track that was present in the
 oligonucleotide that was used to prime the synthesis of first
 strand cDNA and therefore this may represent a bonafide poly A
 tail. The sequence tag present in the cDNA between the NotI site
 and the oligo-dT track served to identify it as a clone from the
 normalized rat eye library cDNA library preparation M.P. Soares
 Lab Clone distribution: clones will be available through Research
 Genetics (www.resgen.com) The following repetitive elements were
 found in this cDNA sequence: 1-44, >POLY_A#Simple_repeat
 Seq primer: M13 Forward
 POLYA-Yes.

FEATURES
 Source
 Location/Qualifiers
 1..654
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="01-R-CNI-cj-f-d-03-0-01"
 /clone_lib="01-R-CNI"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pT7130-Pac (Pharmacia) with a modified
 polylinker. Site_1: Not I; Site_2: Eco RI; The Cj-F CNI
 library is a subtracted library derived from the following
 pool of seven normalized rat libraries: normalized rat
 seminal vesicles, normalized rat penis, normalized rat
 bladder, normalized rat cervix, normalized rat brown
 adipose, normalized rat lungs, and normalized rat
 salivary gland. It was constructed according to the
 procedure described by Bonaldo, Lennon & Soares (Genome
 Research Genome 6: 791-806, 1996). For construction of
 the CNI library, plasmid DNA from the pool of normalized
 libraries was electroporated into competent bacteria for
 the production of single-stranded circular DNA. This was
 then used as a tracer in a subtractive hybridization with
 a driver (PCR amplified inserts from a plasmid DNA template
 preparation) comprising: a) a pool of about 34,000 clones
 from the Rat Unigene Set corresponding to plates R-5-AA-NN
 excluding plates R-5-MM and MN. This pool represented 408
 of the final driver population. b) a pool of about 29,000
 clones from subtracted libraries CA0 and CA1 corresponding
 to plates R-CA0-AMW through R-CA0-BHY, R-CA0-AZX through
 R-CA0-BE, R-CA0-BKE through R-CA0-BKJ, R-CA0-BKP through
 R-CA0-BKS, R-CA0-BKV, R-CA0-BLX through R-CA0-BM, R-
 CA0-BMC through R-CA0-BME, R-CA0-BNS, R-CA0-BOB through
 R-CA0-BOD, R-CA0-BPD through R-CA0-BPG, R-CA1-BBA through
 R-CA1-BDA, R-CA1-BPA through R-CA1-BJF, R-CA1-BJR,
 R-CA1-BJT through R-CA1-BKB, R-CA1-BKD, R-CA1-BKF,

R-CA1-BKI, R-CA1-BKT, R-CA1-BLF, R-CA1-BLI, R-
 CA1-BLN, R-CA1-BLS, R-CA1-BLV, R-CA1-BNR,
 R-CA1-BPF, R-CA1-BPG, R-CA1-BPH, R-CA1-BPI,
 R-CA1-BPJ, R-CA1-BPK, R-CA1-BPL, R-CA1-BPM,
 R-CA1-BPN, R-CA1-BPO, R-CA1-BPP, R-CA1-BPQ,
 R-CA1-BPR, R-CA1-BPS, R-CA1-BPT, R-CA1-BPU,
 R-CA1-BPV, R-CA1-BPW, R-CA1-BPX, R-CA1-BPY,
 R-CA1-BPZ, R-CA1-BQA, R-CA1-BQB, R-CA1-BQC,
 R-CA1-BQD, R-CA1-BQE, R-CA1-BQF, R-CA1-BQG,
 R-CA1-BQH, R-CA1-BQI, R-CA1-BQJ, R-CA1-BQK,
 R-CA1-BQL, R-CA1-BQM, R-CA1-BQN, R-CA1-BQO,
 R-CA1-BQP, R-CA1-BQQ, R-CA1-BQR, R-CA1-BQS,
 R-CA1-BQT, R-CA1-BQU, R-CA1-BQV, R-CA1-BQW,
 R-CA1-BQX, R-CA1-BQY, R-CA1-BQZ, R-CA1-BRA,
 R-CA1-BRB, R-CA1-BRC, R-CA1-BRD, R-CA1-BRE,
 R-CA1-BRF, R-CA1-BRG, R-CA1-BRH, R-CA1-BRI,
 R-CA1-BRJ, R-CA1-BRK, R-CA1-BRL, R-CA1-BRM,
 R-CA1-BRN, R-CA1-BRO, R-CA1-BRP, R-CA1-BRQ,
 R-CA1-BRR, R-CA1-BRS, R-CA1-BRT, R-CA1-BRU,
 R-CA1-BRV, R-CA1-BRW, R-CA1-BRX, R-CA1-BRY,
 R-CA1-BRZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO, R-CA1-BSQ, R-CA1-BSR, R-CA1-BSU,
 R-CA1-BSV, R-CA1-BSW, R-CA1-BSX, R-CA1-BSY,
 R-CA1-BSZ, R-CA1-BSA, R-CA1-BSB, R-CA1-BSF,
 R-CA1-BSG, R-CA1-BSH, R-CA1-BSI, R-CA1-BSJ,
 R-CA1-BSK, R-CA1-BSL, R-CA1-BSM, R-CA1-BSN,
 R-CA1-BSO,



CC tolerance, for immunotherapy, for immunosuppression, for the
 CC treatment of autoimmune diseases, the treatment of allergy, and/or
 CC for inhibiting HIV infection (claimed). The compound that binds to
 CC the C-type lectin is chosen from mannose, fucose, plant lectins,
 CC antibiotics, sugars, proteins and antibodies against C-type
 CC lectins, especially against DC-SIGN. Anti-DC-SIGN antibodies were
 CC developed that not only inhibited transient DC-T-cell interactions
 CC and DC-induced T-cell proliferation but also inhibited HIV-1
 CC infection of DCs.

CC Sequence 404 AA:

Query Match 83.3% Score 1412.5; DB 21, Length 404;
 Best Local Similarity 71.9%; Pred. No. 5,66-107;
 Matches 282; Conservative 20; Mismatches 19; Indels 71; Gaps 4;

QY 1 MSDSKEPRVQQLGL-----OCIGHGALVQLLSFPMILASVLA1 40
 DB 1 MSDSKEPRVQQLGLLEEQRLGLGFRUTRGYKSLAGCLGHPLVQLLSFTLLAG----L 56
 QY 41 LVQVSKVPSLSIQSESDAAYONITQLKAAGELSEKSLQEIYQELTQLKAAGELPE 100
 DB 57 LVQVSKVPSLSIQSESDAAYONITQLKAAGELSEKSLQEIYQELTQLKAAGELPE 116
 QY 101 KSKLQEIYQELTQLKAAGELPEKSKL-----L27
 DB 117 KSKLQEIYQELTQLKAAGELPEKSKLQEIYQELTQLKAAGELPEKSKMQEIYQELTRL 176
 QY 128 -----QEIYQELTQLKAAGELPEKSKLQEIYQELTQLKAAGELPEKSKQ 174
 DB 177 KAAGELPEKSKQEIYQELTQLKAAGELPEKSKQEIYQELTQLKAAGELPEKSKQ 236
 QY 175 QIYQELTDLKTAHERLCPHPCFDMTFPGNCFYMSNSGRNMHDSVATACQEVRAQIVV1 244
 DB 237 ELYQELTQLKAAGELCHCPWEMTFPGNCFYMSNSGRNMHDSVATACQEVRAQIVV1 296
 QY 235 AEEQNFLOLQTSRSNRFESMGLSDLNQPGTQWQVQSPISFQPYWNSGCEPNNSQNEQC 294
 DB 297 AEEQNFLOLQTSRSNRFESMGLSDLNQPGTQWQVQSPISFQPYWNSGCEPNNSQNEQC 355
 QY 295 AEESSGWNNDKQNDVNWICKPAA-CFEPDE 325
 DB 457 AEESSGWNNDKQNDVNWICKPAA-CFEPDE 388

RESULT 3

AAB28614 ID AAB28614 standard; Protein: 404 AA.

XX AAB28614;

XX 12-FEB-2001 (first entry)

XX Human C-type lectin receptor.

XX Human: immunomodulatory; antiallergic; anti-HIV; antiinflammatory;
 XX antidiabetic; antihypertoid; antirheumatic; antiarthritic; vaccine;
 XX C-type lectin receptor; immunotherapy; immunosuppression;
 XX transplant rejection; autoimmune disease; thyroiditis;
 XX rheumatoid arthritis; multiple sclerosis; autoimmune diabetes;
 XX systemic lupus erythematosus; HIV infection; allergy.

XX Homo sapiens.

XX W020006.1251-A1.

XX 26-OCT-2000.

XX 19-APR-2000; 2000W0-NL00253.

XX 19-APR-1999; 99EP-0201204.

XX 20-JAN-2000; 200005-0176924.

PA (UNIV-) UNIV NIJMEGEN.
 PI Figdor CG, Gelfandbeek TH, Van Kooyk Y, Forstma R
 XX WPI: 2000-656424/63.
 XX N-PSDB: AAC65383.
 XX Methods for modulating dendritic cell and T cell interaction;
 PT compound, such as mannose carbohydrates, that binds to a C-type
 PT on a dendritic cell surface, useful in the treatment of auto-
 PT diseases and allergy -
 PS Disclosure: Fig 9; 66pp; English.
 CC The present sequence as DC-SIGN, a HLA-C type lectin receptor
 CC methods for modulating dendritic cell and T cell interaction;
 CC disclosed. The methods involve using a compound that binds to
 CC lectin on the surface of a dendritic cell. The methods are use-
 CC preventing or inhibiting immune responses to specific antigens
 CC inducing tolerance, for immunotherapy, for immunosuppression (e.g.
 CC preventing transplant rejection), for the treatment of autoim-
 CC diseases (e.g., thyroiditis, rheumatoid arthritis, multiple sys-
 CC autoimmune diabetes, systemic lupus erythematosus), HIV infec-
 CC and/or for the treatment of allergy. Antibodies are useful for
 CC detecting the presence of dendritic cells in a biological sam-
 CC or their fragments or epitopes in a biological sample.
 CC XX
 XX Sequence 404 AA:
 Query Match 83.3% Score 1412.5; DB 21, Length 404;
 Best Local Similarity 71.9%; Pred. No. 5,66-107;
 Matches 282; Conservative 20; Mismatches 19; Indels 71;
 QY 1 MSDSKEPRVQQLGL-----OCIGHGALVQLLSFPMILASV1 40
 DB 1 MSDSKEPRVQQLGLLEEQRLGLGFRUTRGYKSLAGCLGHPLVQLLSFTLLAG----L 56
 QY 41 LVQVSKVPSLSIQSESDAAYONITQLKAAGELSEKSLQEIYQELTQLKAAGELPE 100
 DB 57 LVQVSKVPSLSIQSESDAAYONITQLKAAGELSEKSLQEIYQELTQLKAAGELPE 116
 QY 101 KSKLQEIYQELTQLKAAGELPEKSKL-----L27
 DB 117 KSKLQEIYQELTQLKAAGELPEKSKLQEIYQELTQLKAAGELPEKSKMQEIYQELTRL 176
 QY 128 -----QEIYQELTQLKAAGELPEKSKLQEIYQELTQLKAAGELPEKSKQ 174
 DB 177 KAAGELPEKSKQEIYQELTQLKAAGELPEKSKQEIYQELTQLKAAGELPEKSKQ 236
 QY 175 QIYQELTDLKTAHERLCPHPCFDMTFPGNCFYMSNSGRNMHDSVATACQEVRAQIVV1 244
 DB 237 ELYQELTQLKAAGELCHCPWEMTFPGNCFYMSNSGRNMHDSVATACQEVRAQIVV1 296
 QY 235 AEEQNFLOLQTSRSNRFESMGLSDLNQPGTQWQVQSPISFQPYWNSGCEPNNSQNEQC 294
 DB 297 AEEQNFLOLQTSRSNRFESMGLSDLNQPGTQWQVQSPISFQPYWNSGCEPNNSQNEQC 355
 QY 295 AEESSGWNNDKQNDVNWICKPAA-CFEPDE 325
 DB 457 AEESSGWNNDKQNDVNWICKPAA-CFEPDE 388

RESULT 4

AAG79086 ID AAG79086 standard; Protein: 404 AA.

XX AAG79086;

XX 10-DEC-2001 (first entry)

XX Human DC-SIGN, a dendritic cell-specific C-type lectin.

CC human liver cDNA clone HP01347 (see AAV84361). The encoded protein
CC is characterized as a type-II membrane protein having a single
CC N-terminal transmembrane domain. It has 85 kD homology over 294
CC amino acid residues to the human HIV envelope glycoprotein gp120-
CC binding C-type lectin (a CD4-independent HIV receptor) and may
CC function as a receptor on the membrane surface. The invention
CC provides nucleotide sequences (see AAV84359-76) coding for 18
CC transmembrane proteins (see AAV84351-58), vectors containing such
CC polynucleotides, and eukaryotic cells containing the vectors. The
CC proteins can be used as antigens or as compositions in the
CC preparation of antibodies against the proteins. The polynucleotides
CC can be used as probes for gene diagnosis, and as gene sources for
CC gene therapy and anti-sense production of proteins encoded by the
CC cDNA. The host cells are used for the detection of ligands
CC corresponding to the expressed proteins, and the screening of low
CC mol.wt. medicines.

CC Sequence 296 AA:

Query Match 68.2%; Score 1157; DB 20; Length 296;

Best Local Similarity: 83.8%; Freq. No. 2,666; Indels 46; Gaps 1;
Matches 238; Conservative 0; Mismatches 0;

DB 1 MSISKEPVVQGLGLGCLGAGALVQLSPMLAGVAVAIIVGVSKVPSLSQPSQDA 60
DB 1 MSDSKRPVQGLGLGCLGAGALVQLSPMLAGVAVAIIVGVSKVPSLSQPSQDA 60
DB 1 YQNTLTOLKAAGVLESEKSKLOEIVQELTOLKAAGVLESEKSKLOEIVQELTOLKAAGV 120
DB 61 YQNTLTOLKAAGVLESEKSKLOEIVQELTOLKAAGVLESEKSKLOEIVQELTOLKAAGV 120
DB 61 YQNTLTOLKAAGVLESEKSKLOEIVQELTOLKAAGVLESEKSKLOEIVQELTOLKAAGV 120
DB 121 LPEKSKLOEIVQELTOLKAAGVLESEKSKLOEIVQELTOLKAAGVLESEKSKLOEIVQEL 159
DB 121 LPEKSKLOEIVQELTOLKAAGVLESEKSKLOEIVQELTOLKAAGVLESEKSKLOEIVQEL 180
DB 160 -----LKAAGVLEPQSKQOQIVQELTDLTKTAFERLCRHC 194
DB 181 TELKAAGVLEPQSKLOEIVQELTOLKAAGVLEPQSKQOQIVQELTDLTKTAFERLCRHC 240
DB 195 PKDWTFFQNCYFMSNSQNRHDSVTACQEVRAQLVVIKTAERQ 238
DB 241 PKDWTFFQNCYFMSNSQNRHDSVTACQEVRAQLVVIKTAERQ 284

RESULT 8

AA079093 ID AAG79093 standard; Protein: 296 AA.

AA079093 AC AAG79093;

UT 10-DEC-2001 (first entry)

DE Amino acid sequence of a splice variant of human DC-SIGN.

KW Human; receptor: DC-SIGN; dendritic cell; T lymphocyte; HIV,

KM gp120; C-type lectin; ICAM3; HIV entry; T cell; macrophage;

KW HIV infection; splice variant.

OS Homo sapiens.

PN W0200164752-A2.

PD 07-SEP-2001.

PE 28-11-2001, 2001WO-0506322.

PR 02-MAR-2000; 2000US-0517605.

PA (UNIV.) UNIV NEW YORK STATE

XX (UNIV.) UNIV NIDMEGEN.

XX LITMAN DR, KWON D, VAN KOOYK Y, GELJENBECK T;

DR WPI: 2001-602565/68.

XX N-PSDB: AA165469.

PT An antibody for the treatment of prevention of HIV-infection comprises

XX a gp120 portion which binds to DC-SIGN or is exposed upon gp120 binding

XX or DC-SIGN due to concomitant conformational change.

XX Example 1; Page 130-131; 131pp; English.

CC The present sequence represents a splice variant of a human receptor
CC designated DC-SIGN. DC-SIGN is specifically expressed on dendritic cells
CC and facilitates infection of T lymphocytes with HIV. DC-SIGN is identical
CC to a HIV-1 gp120-binding C-type lectin. DC-SIGN binds ICAM-3 (which is
CC expressed constitutively on T lymphocytes) with high affinity. The
CC specificity of gp120. This antigenic fragment binds to DC-SIGN or is exposed
CC upon gp120 binding of DC-SIGN due to concomitant conformational change.
CC The antibody inhibits the trans enhancement of HIV entry into a T
CC cell or macrophage facilitated by dendritic cells. The antibody is
CC useful to treat or prevent HIV infection.

CC Sequence 296 AA:

Query Match 68.2%; Score 1157; DB 22; Length 296;

Best Local Similarity: 83.8%; Freq. No. 2,666; Indels 46; Gaps 1;
Matches 238; Conservative 0; Mismatches 0;

DB 1 MSISKEPVVQGLGLGCLGAGALVQLSPMLAGVAVAIIVGVSKVPSLSQPSQDA 60
DB 1 MSDSKRPVQGLGLGCLGAGALVQLSPMLAGVAVAIIVGVSKVPSLSQPSQDA 60
DB 1 YQNTLTOLKAAGVLESEKSKLOEIVQELTOLKAAGVLESEKSKLOEIVQELTOLKAAGV 120
DB 61 YQNTLTOLKAAGVLESEKSKLOEIVQELTOLKAAGVLESEKSKLOEIVQELTOLKAAGV 120
DB 61 YQNTLTOLKAAGVLESEKSKLOEIVQELTOLKAAGVLESEKSKLOEIVQELTOLKAAGV 120
DB 121 LPEKSKLOEIVQELTOLKAAGVLESEKSKLOEIVQELTOLKAAGVLESEKSKLOEIVQEL 159
DB 121 LPEKSKLOEIVQELTOLKAAGVLESEKSKLOEIVQELTOLKAAGVLESEKSKLOEIVQEL 180
DB 160 -----LKAAGVLEPQSKQOQIVQELTDLTKTAFERLCRHC 194
DB 181 TELKAAGVLEPQSKLOEIVQELTOLKAAGVLEPQSKQOQIVQELTDLTKTAFERLCRHC 240
DB 195 PKDWTFFQNCYFMSNSQNRHDSVTACQEVRAQLVVIKTAERQ 238
DB 241 PKDWTFFQNCYFMSNSQNRHDSVTACQEVRAQLVVIKTAERQ 284

RESULT 9

ABR09715 ID ABR09715 standard; Protein: 296 AA.

AC ABR09715;

UT 11 JUN-2002 (first entry)

DE Amino acid sequence of human polypeptide HP01347.

KW Antibody; antigen; transmembrane domain protein; HP01347.

OS Homo sapiens.

PN W0200208415-A1.

PD 31-JAN-2002.

PE 24-JUL-2001; 2001WO-JP06371.

PR 24-JUL-2000; 2000JP-0222773.

PA 24-ACQ 2000; 2000JP-0254407.

XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.

PI Kato S, Nagata N, Fujimura N, Kobayashi M, Ito K, Ishizuka Y
XX WPI: 2002-195877/25.
DR N-PSDB: ABL1993.

PT Antibody preparation by inoculation of an animal with a vector
XX expressing a fusion protein of an antigen on the C-terminal side of a
PT transmembrane domain for use as drugs, diagnostic reagents and
PT laboratory reagents

XX Example: Page 23-27; 45pp; Japanese

XX The specification describes a method of antibody preparation. The
CC method comprises inoculating an animal with a vector expressing
CC a fusion protein having an antigen protein fused to the C-terminal
CC side (extracellular) of a transmembrane domain protein (the
CC N-terminal side of which is intracellular), and then isolating and
CC purifying the antibody from the animal. The antibodies can be used
CC as drugs, vaccines, diagnostic reagents and laboratory reagents. The
CC present sequence represents a polypeptide, designated HP01347, which was
CC used in the course of the invention.

XX Sequence 296 AA:

Query Match 68.2% Score 1157; DR 33; Length 296;
Best Local Similarity 83.8%; Pred. No. 2; 6e-86;
Matches 238; Conservative 0; Mismatches 0; Indels 46; Gaps 1;

OY 1 MSDSKPRVQGLGGLGCTHNAIVIGLSPMLLAGYVAIIVGVSKVPSLSQESQDA 60

DB 1 MSDSKPRVQGLGGLGCTHNAIVIGLSPMLLAGYVAIIVGVSKVPSLSQESQDA 60

OY 61 IVNLTQLAAVSELSESKLQEIYQELTQLAAVSELSESKLQEIYQELTQLAAV 20

DB 61 IVNLTQLAAVSELSESKLQEIYQELTQLAAVSELSESKLQEIYQELTQLAAV 20

OY 121 LPEKSKLQEIYQELTQLAAVSELSESKLQEIYQELTQLAAVSELSESKLQEIY 159

DB 121 LPEKSKLQEIYQELTQLAAVSELSESKLQEIYQELTQLAAVSELSESKLQEIY 159

OY 160 -----LKAAGELPDQSKQVQIYQELTQLAAVSELSESKLQEIYQELT 194

DB 161 LELKAAGELPEKSKLQEIYQELTQLAAVSELSESKLQEIYQELTQLAAV 240

OY 195 PKQWTFQGCYFMSNSQPMHDSVTACQEPFAOTVYTKFAEFO 238

DB 241 PKQWTFQGCYFMSNSQPMHDSVTACQEPFAOTVYTKFAEFO 284

RESULT 10
A0088024
ID A0088024 standard; Protein: 325 AA.

XX A0088024:

DT 05-JUN-2002 (first entry)

XX Mouse protein encoded by Contig 1A.

XX Nucleic acid library; immune response, asthma, COPD;

XX airway hyperresponsiveness; bronchoalveolar manifestation;

XX signature sequence; SS; chronic obstructive pulmonary disease;

XX allergic disease; rhinitis; atopic dermatitis; urticaria;

XX autoimmune disease; multiple sclerosis; inflammatory bowel disease;

XX allograft rejection; infectious disease.

XX calcium-activated chloride channel.

XX Mus sp.

XX W0200214366-A2.

XX 21-FEB-2002.

XX

XX

XX

PF 16-AUG-2001; 2001W0-NL00610
XX
XX 16-AUG-2001; 2001W0-NL00610
XX
XX (OYUT-) RIJSDUNIV UTRECHT.

XX Grool PC, Van Berghencouwen HJ, Van Oosterhout AJM;

XX WPI, 2002; 241888/29.

XX N-PSDB: ABK47192.

XX Nucleic acid library comprising genes which are capable of

XX progression and suppression of an immune response, especially

XX response observed with airway hyperresponsiveness of asthma

XX Example 10; Fig 10; 120pp; English.

XX The invention relates to a nucleic acid library comprising ge

XX their fragments which are capable of modulating an immune res

XX observed with airway hyperresponsiveness and/or bronchial

XX manifestations of asthma. Also included are a method for modu

XX immune response of an individual comprising modulating a gene

XX a nucleic acid at least functionally equivalent to a nucleic

XX as R1-S0-R1-A11, S001-A10, S002-1-C11, S001-A12, and R1-S0-R1

XX substance (for use as a medicament) capable of modulating a g

XX acid identifiable by SS and the use of a proteinaceous subst

XX from a nucleic acid at least functionally equivalent to a nu

XX medicament) against the substance. The antagonist (for a

XX useful for the treatment of an immune response observed with

XX hyperresponsiveness and/or bronchoalveolar manifestations of

XX The method is useful for modulating the above immune respo

XX gene encodes a gene product capable of modulating the immune

XX The substance is useful for treating an immune response, fat

XX asthma, chronic obstructive pulmonary disease (COPD), aller

XX (rhinitis atopic dermatitis, urticaria), autoimmune disease

XX multiple sclerosis), inflammatory bowel disease, allograft

XX infectious disease. The present sequence is a mouse or human

XX protein encoded by a signature sequence gene or its homoloqu

XX equivalent.

XX Sequence 325 AA:

Query Match 42.5%; Score 740; DR 23; Length 325

Best Local Similarity 43.0%; Pred. No. 1; 3e-50;

Matches 154; Conservative 49; Mismatches 77; Indels 7;

OY 1 MSDSKPRVQGLG-----LLOGLGHC--ALV 1

DB 1 MSSTKAKMOTLSSMDDELWVSGSPYVTSPIPNSTKCLAGSGSHVPLV 1

OY 30 FMLLAGYVAIIVGVSKVPSLSQESQDAIYQNTLQ-----KAAGELSEK 30

DB 61 FFLAGLLITLFOVSKTPN--TEQKPKERILLOELTQLDELTSKIPISQGRNS 61

OY 84 IYQELTQLAAVSELSESKLQEIYQELTQLAAVSELSESKLQEIYQELTQLAA 84

DB 119 IYQELTQLAAVSELSESKLQEIYQELTQLAAVSELSESKLQEIYQELTQLAA 119

OY 144 LPEKSKLQEIYQELTQLAAVSELSESKLQEIYQELTQLAAVSELSESKLQEIY 144

DB 158 -----ELSKLSPVVDSDSKQKRTIYQELTQLAAVSELSESKLQEIYQELT 158

OY 202 QGNCYFMSNSQPMHDSVTACQEPFAOTVYTKFAEFO 238

DB 203 LGNCYFMSNSQPMHDSVTACQEPFAOTVYTKFAEFO 238

OY 262 EATLWVDSITLSSRFQYVWNRGPNNIGFEDCVFAGDGNDSKCELRK 262

DB 262 EATLWVDSITLSSRFQYVWNRGPNNIGFEDCVFAGDGNDSKCELRK 262

RESULT 11

AA088026 standard: Protein: 325 AA.

AA088026:

05-JUN-2002 (first entry)

Human DC-SIGN protein.

Nucleic acid library: immune response; asthma; COPD;

airway hyperresponsiveness; bronchial hyperresponsiveness;

signature sequence; SS; chronic obstructive pulmonary disease;

allergic diseases; rhinitis; atopic dermatitis; urticaria;

autoimmune disease; multiple sclerosis; inflammatory bowel disease;

allergic rejection; infectious disease;

calcium-activated chloride channel;

Human sapiens;

W:200214366-A2.

21-FEB-2002.

16-AUG-2001: 2001W-NL00610.

16-AUG-2000: 2000P-090967.

(0901-) BLKSNV UTRECHT.

Genot PC: Van Bergenhenegouwen HJ, Van Westerbout AM.

WPI: 2002-241888/29.

Nucleic acid library: asthma; genes which are capable of inducing

protection and suppression of an immune response, especially an immune

response observed with airway hyper-responsiveness of asthma.

Example 10: Fig 13: 129P; English.

The invention relates to a nucleic acid library comprising genes or

their fragments which are capable of modulating an immune response

observed with airway hyperresponsiveness and/or bronchial hyper-

responsiveness of asthma. Also included are a method for modulating an

immune response of an individual comprising modulating a gene comprising

a nucleic acid at least functionally equivalent to a nucleic acid

identifiable by a signature sequence (SS) given in the specification such

as P1-SS-P1-A11, SS-P1-A10, SS-P1-A11, SS-P1-A12, and P1-SS-P1-A12, a

substance (for use as a modulator) capable of modulating a gene

comprising a nucleic acid at least functionally equivalent to a nucleic

acid identifiable by SS and the use of a proteinaceous substance derived

from a nucleic acid at least functionally equivalent to a nucleic acid

identifiable by SS for the production of an antagonist (for use as a

modulator) against the substance. The antagonist and substance are

useful for the treatment of an immune response observed with airway

hyperresponsiveness and/or bronchial hyperresponsiveness of asthma.

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

1 MSUSKPPWQVQDQ-----LIGGCGHC- ALVIGLIS 29

Tue Dec 10 10:10:07 2002

us-09-831-458a-12.std.rag

Page 11

OY 315 CKKPAA 320
11:11:11
DB 183 GEOPSA 188

Search Completed: December 7, 2002, 11:03:57
Job time : 76 secs

•
•
•
•

CC novel ICAM-3 receptor that is specifically expressed by human DC.
CC It is involved in the initial transient DC-1-cell interaction
CC necessary for initiating an immune response. DC-SIGN is also able
CC to bind to the HIV envelope protein gp120 and to facilitate HIV-1
CC entry into DC. The cDNA was obtained by RT-PCR amplification (see
CC AAAB8741.1-42) of DC RNA. The invention relates to the use of a
CC compound that binds to a C-type lectin (especially DC-SIGN) on the
CC surface of a DC, and in the preparation of a composition for
CC modulating, especially reducing, the immune response in an animal,
CC in particular a human or other mammal. The composition modulates
CC the interactions between a DC and a T-cell, more specifically
CC between a C-type lectin on the surface of a DC and an ICAM receptor
CC on the surface of a T-cell. It is used to prevent/inhibit immune
CC responses to specific antigens, for inducing tolerance, for
CC immunotherapy, for immunosuppression, for the treatment of
CC autoimmune diseases, the treatment of allergy, and/or for
CC inhibiting HIV infection (claimed). The compound that binds to
CC the C-type lectin is chosen from mannose, fucose, plant lectins,
CC antibiotics, sugars, proteins and antibodies against C-type lectins.
XX
50 Sequence 1215 BP; 315 A, 321 C, 349 G, 230 T; 0 other;

Sequence 1215 BP; 315 A, 321 C, 349 G, 230 T; 0 others.

Query Match	43.26	Score	665.4	E ₀	51	Length	1215
Best Local Similarity	88.83	Prod. No.	7.6e+173				
Matches	720	Conservative	0	Mismatches	51	Indels	0
						Gaps	0

QY	424	TGGGAGCAATCGGACAGGACGACATCTTACGAGAACTGGACGACATTAACCTGAT	483
Db	342	TCCAGAGAAATCTTAAGCTGACAGAAATCTACGAGAAAGCTGGACGCTGTCAAT	401
QY	484	GGGTGAGCTCTCAGAGCAATTCACAGCTGCAGAGATCTACGAGAACTGACCTGAA	543
Db	402	GGGTGAGCTTTCAGAGCAATTCAGAGATCTACGAGAACTGACCTGACCTGAA	461
QY	544	GGGTGAGTGGGTGAGCTTTCAGAGCAATTCAGAGATCTACGAGAACTGACCTGAC	603
Db	462	GGGTGAGTGGGTGAGCTTTCAGAGCAATTCAGAGATCTACGAGAACTGACCTGAC	521
QY	604	GGGTGAGTGGGTGAGCTTTCAGAGCAATTCAGAGATCTACGAGAACTGACCTGAC	663
Db	522	GGGTGAGTGGGTGAGCTTTCAGAGCAATTCAGAGATCTACGAGAACTGACCTGAC	581
QY	664	GGGTGAGTGGGTGAGCTTTCAGAGCAATTCAGAGATCTACGAGAACTGACCTGAC	723
Db	582	GGGTGAGTGGGTGAGCTTTCAGAGCAATTCAGAGATCTACGAGAACTGACCTGAC	641
QY	724	GGGTGAGTGGGTGAGCTTTCAGAGCAATTCAGAGATCTACGAGAACTGACCTGAC	783
Db	642	GGGTGAGTGGGTGAGCTTTCAGAGCAATTCAGAGATCTACGAGAACTGACCTGAC	701
QY	784	GGGTGAGTGGGTGAGCTTTCAGAGCAATTCAGAGATCTACGAGAACTGACCTGAC	843
Db	702	GGGTGAGTGGGTGAGCTTTCAGAGCAATTCAGAGATCTACGAGAACTGACCTGAC	761
QY	844	GGGTGAGTGGGTGAGCTTTCAGAGCAATTCAGAGATCTACGAGAACTGACCTGAC	903
Db	762	GGGTGAGTGGGTGAGCTTTCAGAGCAATTCAGAGATCTACGAGAACTGACCTGAC	821
QY	904	GGGTGAGTGGGTGAGCTTTCAGAGCAATTCAGAGATCTACGAGAACTGACCTGAC	963
Db	822	GGGTGAGTGGGTGAGCTTTCAGAGCAATTCAGAGATCTACGAGAACTGACCTGAC	881
QY	964	GGGTGAGTGGGTGAGCTTTCAGAGCAATTCAGAGATCTACGAGAACTGACCTGAC	1023
Db	882	GGGTGAGTGGGTGAGCTTTCAGAGCAATTCAGAGATCTACGAGAACTGACCTGAC	941
QY	1024	GGGTGAGTGGGTGAGCTTTCAGAGCAATTCAGAGATCTACGAGAACTGACCTGAC	1083
Db	942	GGGTGAGTGGGTGAGCTTTCAGAGCAATTCAGAGATCTACGAGAACTGACCTGAC	1001
QY	1084	GGGTGAGTGGGTGAGCTTTCAGAGCAATTCAGAGATCTACGAGAACTGACCTGAC	1143
Db	1002	GGGTGAGTGGGTGAGCTTTCAGAGCAATTCAGAGATCTACGAGAACTGACCTGAC	1061

[illegible]

RESULT 7

ID	AAC65383	standard, cDNA, 1215 bp

AAC65383;

12-1118-2001 (L115C C01177)

XX

KW antidiabetic; antihyroid; antineumatic; antiarthritic; vaso-

1 rheumatoid arthritis; multiple sclerosis; autoimmune diabetes
2 HIV infection; allergy; DC-SIGN

XX Homo sapiens.

PN WO2000063251-A1

PD 26-OCT-2000

PE 19-APR-2000; 2000MC-NL00253

PR 19-APR-1999; 99EP-0201204
PR 20-JAN-2000; 2000US-0175924

PA (UYNI-) UNIV NIJMEGEN.

PI Fyodor GG, Gekhtenbe

DE WT: 2000-656424/b3.
DE F-PCDB: AAP28614

Methods for modulation

PT on a dendritic ce

[illegible]

The proposed sequence encodes for a 44kDa C-terminal protein.

disclosed. The method

CC preventing or inhibiting immune responses to specific antigens

CC or epitopes in a biological sample.

sequência 1213 BR, 312 M, 322 ...

Query Match	43.28;	Score 665.4;	PR 21;	Length 14.
Best Local Similarity	88.88;	Prod. No. 7.6e-173;		
Matches 720;	Conservative	0;	Mismatches	91;
			Indels	

424 TCAGGAACTATCTGATTAAGATTTAAATTAATAATTTGATCTAGCTTAAAT

Genome version 5.1.3
Copyright (c) 1993 - 2002 CompuLink Ltd

OM protein - protein search, using SW model

Run on: December 7, 2002, 10:41:04 : Search time 44 seconds
(without alignments)
710,084 Million cell updates/sec

Title: US-09-831-458A-12
Perfect score: 166
Sequence: 1 MSDSKPPVQGLLDDQLCH RGVDRWVTKKVAAGHDE 325

Scoring table: RUSUM62
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 9613422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: 1: PIR.73: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1412.5	83.3	404	2	A46274 HIV gp120-binding
2	351	20.7	309	1	IgE Fc receptor II
3	350	20.6	331	1	IgE Fc receptor, I
4	326.5	19.3	207	1	hepatic lectin - C
5	322.5	19.0	321	1	IgE Fc receptor II
6	309.5	18.2	550	2	Kupffer cell recep
7	307	18.1	304	2	lectin, galactose/
8	306	18.0	284	2	asialoglycoprotein
9	289	17.0	284	1	hepatic lectin - r
10	287	16.9	306	2	lectin M-ASGP-BP p
11	271.5	16.0	291	1	hepatic lectin HI
12	267	15.7	311	1	asialoglycoprotein
13	258	15.2	301	2	asialoglycoprotein
14	243	14.3	301	2	hepatic lectin 2 -
15	232	13.7	237	2	type II lectin-11k
16	230.5	13.6	363	2	lectin-like oxidiz
17	230	13.6	742	2	scavenger receptor
18	227.5	13.4	166	2	C-type lectin, B 1
19	219.5	12.9	162	1	lectin BRA3-2 prec
20	218.5	12.9	162	1	lectin BRA3-1 prec
21	216.5	12.8	216	2	natural killer cel
22	210.5	12.4	1456	1	mannose receptor p
23	203.5	12.0	330	2	brevican - human
24	202.5	11.9	323	1	agrecan precursor
25	201.5	11.9	2415	1	agrecan precursor
26	201.5	11.8	2124	2	mannose receptor,
27	199.5	11.7	1479	2	mannose receptor,
28	199	11.7	883	2	brevican precursor
29	198	11.7	257	2	gene 17.5 protein

30	197	11.6	2132	1	A55182	agrecan
31	195	11.5	883	2	S49126	brevican
32	195	11.5	2109	1	I50421	agrecan
33	194.5	11.5	912	2	A54423	brevican
34	193.5	11.4	1340	2	A39608	proteog
35	193.5	11.4	2327	2	T42630	agrecan
36	191	11.3	199	2	JH0822	lymphoc
37	189.5	11.2	1268	2	S52781	neutral
38	188	11.1	237	2	A44467	natural
39	188	11.1	1455	1	A48925	mannos
40	186.5	11.0	1257	1	S28764	neuram
41	180.5	10.6	147	2	A26697	echinov
42	178	10.5	244	1	LNMSMC	mannose
43	176.5	10.4	3562	2	A47171	chondro
44	175	10.3	412	2	R42755	E-type
45	174.5	10.3	391	2	T34284	hypertro

ALIGNMENTS

RESULT 1

A46274 HIV gp120-binding C-type lectin - human

C/Species: Homo sapiens (man)

C/Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 01

C/Accession: A46274

C/Curtis, B.M.; Scharnowski, S.; Watson, A.J.

Proc. Natl. Acad. Sci. U.S.A. 89, 8156-8160, 1992

A/Title: Sequence and expression of a membrane-associated C-type

A/Reference number: A46274; MIM:92490446; PMID:1518865

A/Accession: A46274

A/Status: preliminary

A/Molecule type: nucleic acid

A/Residues: 1-404 <CYP>

A/Experimental source: Placenta

A/Note: sequence extracted from MCB: h-12-10 (NCBI:113134, NCBI)

C/Superfamily: C-type lectin homology

F:256-377/Domain: C-type lectin homology <LCH>

Query Match 83.3% Score 1412.5; DB 2; Length 404

Best Local Similarity 71.9%; Pred. No. 1,20-78;

Matches 282; Conservative 20; Mismatches 19; Indels

QY	1	MSDSKPPVQGLL	-----GIGGALVQLSPMLAT
DB	1	MSDSKPPVQGLL	-----GIGGALVQLSPMLAT
QY	41	LVQYSKPPSSLSQSEQDAIYQNTQI	KAAGVLSKSKLQEIYQELTQKAAV
DB	57	LVQYSKPPSSLSQSEQDAIYQNTQI	KAAGVLSKSKLQEIYQELTQKAAV
QY	101	KSRLQEIYQELTRKAAGVLSKSKL	-----
DB	117	KSRLQEIYQELTRKAAGVLSKSKL	-----
QY	128	-----EFYQELTFYAAVTFYPPSKY	QFYQHTYTK KAAGVLSK
DB	177	KAAGVLSKSKLQEIYQELTRKAAGVLSKSKLQEIYQELTRKAAGVLSK	-----
QY	175	QIYQELTQKAAVLSKSKLQEIYQELTRKAAGVLSKSKLQEIYQELTRKAAGVLSK	-----
DB	237	EIYQELTQKAAVLSKSKLQEIYQELTRKAAGVLSKSKLQEIYQELTRKAAGVLSK	-----
QY	235	AEQNFQLOLSPPSNFPMGSLSDI	NOFTQWQVSGSLSPFQRYNNSGPRNS
DB	237	AEQNFQLOLSPPSNFPMGSLSDI	NOFTQWQVSGSLSPFQRYNNSGPRNS
QY	295	AEFSQSMKNGVQVNTYICKKVA	CPYP 725
DB	357	AEFSQSMKNGVQVNTYICKKVA	CPYP 725

F146-321/Domain: extracellular #status: predicted -EXT-
 F164-84/Region: 21-residue repeat
 F181-321/Region: 21-residue repeat
 F185-105/Region: 21-residue repeat
 F106-126/Region: 21-residue repeat
 F106-126/Region: 21-residue repeat
 F125-321/Product: soluble 146-binding factor (29%) #status: predicted -136
 F148-321/Product: soluble 146-binding factor (25-27%) #status: predicted -136
 F150-321/Product: soluble 146-binding factor (25-27%) #status: predicted -136
 F163-282/Domain: C-type lectin homology -LCH-
 F163-282/Domain: C-type lectin homology -LCH-
 F147-148/Cleavage site: Lys-Leu (unidentified proteinase) #status: experimental
 F149-150/Cleavage site: Arg Met (unidentified proteinase) #status: experimental
 F191-262/259-273/Disulfide bonds: #status: experimental

Query Match 19.08; Score 322.5; DB 1; Length 321;
 Best local similarity 28.68; Pred. No. 1 to 12;
 Matches 95; Conservative 61; Mismatches 121; Indels 55; Gaps 13;

QY 1 MSDSPKPVQDLGLGLGHA--IVQLDLSMLAGVALLVQVSKVSSLSQ--EUSE 57
 DB 1 MERRQYSEIEIPRRRCRRRTQIVLGLVTANWALITLILIMHTQTSKQLEFPA 60
 QY 58 QDAIVNTDQKAAG-ELSEKSKQPTVETLQIKAAVFLPEKSKIQETLPLKA 116
 DB 61 APNVSVSKNIFSHHRCQMOKSSTQTSLELELA-----EGQELK 134
 QY 117 ANGELERKSLQETLQETLQIKAAVCELPEKSK--LQFVYQFIPLKAAVLELQVSK 172
 DB 105 Q--DLISLNNLNGLDLSSFSKQ--ELNERNASDLLELREEVKLR----- 149
 QY 173 QQQVYELTDLKATFEELPCHTQEWITFQJCYEYNSNENHNSVTAQLVKAQAVV 212
 DB 150 -----MLQVSSGVQVCTCPKKNINFCQKYEGSKQVHAPVMEQGVVSI 201
 QY 233 KTAEGNPLDQGPSNPFSSWMIQINQGTQVQVNGSPVLSFQVYNSGPP--NNSGN 294
 DB 232 HSEFQEFELKIASHTG--SWISLKNLLKGLIWLWAKSHVVS---NMAHGEETSKSQ 236
 QY 292 EHCAPFSQSTWNNPCTVW NWTGKRTAAQ 321
 DB 257 EDCVMHSGRMADAFCDPRKLGAWVDELATG 288

RESULT 6
 A28166
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 05-Nov-1999
 C:Accession: A38674; A28166
 F:History: G.W.: Hill, P.L.
 J: Biol. Chem. 265, 1850-1857, 1991
 A:Title: Structure of the gene for a carbohydrate-binding receptor analog for rat Kupffer
 A:Reference number: A38674; M01D:91107889; PMID:1846367
 A:Accession: A38674
 A:Status: preliminary
 A:Molecule type: RNA
 A:Residues: 1-550 (H02)
 A:Cross-references: GH:M55532; NID:420302; PUBM AAA40892.1; PID:420302
 F:History: G.W.: Hill, R.L.
 J: Biol. Chem. 263, 7487-7492, 1988
 A:Title: Molecular cloning and sequencing of a cDNA for a carbohydrate binding receptor
 A:Reference number: A28166; M01D:98227933; PMID:286487
 C:Accession: A28166
 A:Molecule type: mRNA
 A:Residues: 1-550 (H02)
 A:Cross-references: GH:J03734; NID:420750; PUBM AAA11472.1; PID:420750
 C:Supplement: C-type lectin homology
 C:Keywords: transmembrane protein
 F:412-536/Domain: C-type lectin homology -LCH-

Query Match 18.28; Score 309.5; DB 2; Length 550;
 Best local similarity 28.78; Pred. No. 1 to 11;

Matches 81; Conservative 45; Mismatches 110; Indels 45; Gaps 11;

QY 52 QYVSTPQAVYGN--LGLAAVDEHSEKTEGLTYLQLKAAVCEPEKSKLQGLYOR 110
 DB 286 AGVGNAMH--LGLLGLVFAELKELSEN--SGLVNVNCKLMSSEELQTRK 349
 QY 111 LLEKAAVTLDESKTEGLTYLQLKAAVCEPEKSKLQGLYOR 161
 DB 340 LSVNSAL-----KSNVGMKLSNIGKAAVYSLKTELEKALITAAKLYQVQSLDAOK 394
 QY 162 AAVCHLHNGV--LGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLV 221
 DB 394 AVAHLLGKLTGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLV 449
 QY 222 QYVPAQIVVITAPPGNQLGLVSESNPFSSWMIQINQGTQVQVNGSPVLSFQV 280
 DB 440 QVSGAHIAVTSQPKAPVQITNAVH--WGLDNGHGTGNNVWDVGPVYVSGPP 497
 QY 281 WNSGPPNN--SGN--EHCAPFSQSTWNNPCTVW NWTGKRTAAQ 317
 DB 138 WNSGPPNN--SGN--EHCAPFSQSTWNNPCTVW NWTGKRTAAQ 538

RESULT 7
 JX0209
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 20-Aug-1999
 C:Accession: JX0209; PX0009
 F:History: M.: Kawakami, K.; Osawa, T.; Toyoshima, S.
 J: Biochem. 111, 331-336, 1992
 A:Title: Molecular cloning and expression of cDNA encoding a galactose/N-acetylglact
 A:Reference number: JX0209; M01D:9248022; PMID:1587794
 A:Accession: JX0209
 A:Molecule type: mRNA
 A:Residues: 1-394 (SNT)
 A:Cross-references: SE:536776; NID:443260; PUBM AAB2171.1; PID:449361
 F:History: S.: Sato, M.; Toyoshima, S.; Osawa, T.
 J: Biochem. 104, 600-605, 1988
 A:Title: Purification and characterization of a lectin-like molecule specific for gal
 A:Reference number: PX0009; M01D:8917855; PMID:3241002
 A:Accession: PX0009
 A:Molecule type: protein
 A:Residues: 102-120 (197, X, 149-151 <DAs>
 C:Supplement: hepatic lectin; C-type lectin homology
 C:Keywords: glycoprotein; lectin; macrophage; transmembrane protein
 F:36-61/Domain: transmembrane #status: predicted -TPA>
 F:173-294/Domain: C-type lectin homology -LCH>
 F:74,166/Rinding site: carbohydrate (Asn) (occident) #status: predicted

Query Match 18.18; Score 307; DB 2; Length 304;
 Best local similarity 30.38; Pred. No. 8 to 12;
 Matches 97; Conservative 59; Mismatches 101; Indels 59; Gaps 10;

QY 27 LLSMLAGVVAIVGVVVPSPQSPGFGATVNTQYKAAVFLSEKSLQETLQ 86
 DB 36 LLESL--GLSTLIVVAVVSGNSQ-----LKNQVLEKATLNTSKIKAE--EQ 84
 QY 87 ETLQKAAVTLDESKTEGLTYLQLKAAVCEPEKSKLQGLYOR 144
 DB 85 SL-----LKNQVLEKATLNTSKIKAE--EQ 128
 QY 134 LLESLAGVVAIVGVVVPSPQSPGFGATVNTQYKAAVFLSEKSLQETLQ 197
 DB 129 --EKEF-----LGLLGLVFAELKELSEN--SGLVNVNCKLMSSEELQTRK 176
 QY 198 WTLQKAAVTLDESKTEGLTYLQLKAAVCEPEKSKLQGLYOR 257
 DB 177 WTLQKAAVTLDESKTEGLTYLQLKAAVCEPEKSKLQGLYOR 234
 QY 258 LLESLAGVVAIVGVVVPSPQSPGFGATVNTQYKAAVFLSEKSLQETLQ 310
 DB 235 L--LGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLV 291

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2002, 11:27:01 : Search time 2718 Seconds

(Without alignments)
3479,916 Million cell updates/sec

Title: US-09-831-458a-12

Perfect score: 1696

Sequence: 1 MDSKEPPVQSLHCTICH

KCTVDMWICKFACFRFP 325

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Zgapop 6.0, Zgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 20000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL-frame_plus_p2n -DB-seq-length 0 -DB-max-len 20000000
-O/cn2.1/USPTO.spool/US09831458/unal.0512002.103219.15247/36f-Query.fasta.1.515
-DB-GenEmb1 -OEMT=fastlap -SUFFIX=std.rge -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS-bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=humano cpi -LIST=45
-DOCALLON=200 -THR_SCORE=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OFFEMT=ptc -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=usf09-831-458a-12 -JOB=usf09-831-458a-12 -15247 -RCP=6 -RCP=3
-NO_XLPRY -NO_MMAR -LARGEQUERY -NEG_SCORES=0 -WAIT -LON3145 -DEFVTIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6 -FCGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

GenEmb1:
1: gb_fa:
2: gb_hg:
3: gb_in:
4: gb_om:
5: gb_ov:
6: gb_pa:
7: gb_ph:
8: gb_pl:
9: gb_pr:
10: gb_rv:
11: gb_sts:
12: gb_sy:
13: gb_un:
14: gb_vt:
15: fm_fa:
16: fm_fun:
17: fm_hum:
18: fm_in:
19: fm_mu:
20: fm_om:
21: fm_or:
22: fm_ov:
23: fm_pa:
24: fm_ph:
25: fm_pl:
26: fm_ro:
27: fm_sts:
28: fm_un:

29: em_vt:
30: em_hg_hum:
31: em_hg_iny:
32: em_hg_other:
33: em_hg_mus:
34: em_hg_pin:
35: em_hg_rtd:
36: em_hg_mam:
37: em_hg_vtl:
38: em_sy:
39: em_hggo_hum:
40: em_hggo_mus:
41: em_hggo_other:

Pred. No. is the number of results predicted by chance to be
score greater than or equal to the score of the result being
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Descr
1	1646.5	97.1	1800	9	AF290887	AF290887
2	1639	96.6	1200	9	AY042234	AY042234
3	1628	96.0	1510	9	AF245219	AF245219
4	1536.5	90.6	1648	9	AY042235	AY042235
5	1494.5	88.1	1212	6	AY042238	AY042238
6	1412.5	83.3	1212	6	AX342003	AX342003
7	1412.5	83.3	1215	6	AX054403	AX054403
8	1412.5	83.3	1215	6	AY042221	AY042221
9	1412.5	83.3	1312	6	AX287103	AX287103
10	1412.5	83.3	1312	6	HDMLECTING	HDMLECTING
11	1412.5	83.3	1643	6	AF290886	AF290886
12	1412.5	83.3	1643	6	AB015629	AB015629
13	1407	83.0	1643	9	AF361085	AF361085
14	1407	82.8	1212	6	AF166106	AF166106
15	1404.5	82.5	1212	6	AX342004	AX342004
16	1398.5	82.5	1212	6	AX342004	AX342004
17	1398.5	82.5	1212	6	AX342004	AX342004
18	1398.5	82.5	1212	6	AX342004	AX342004
19	1398.5	82.5	1212	6	AX342004	AX342004
20	1378	81.2	1315	9	AY042229	AY042229
21	1375	81.1	1166	9	AY040319	AY040319
22	1375	80.7	1197	9	AY042227	AY042227
23	1369.5	80.5	1146	9	AF343727	AF343727
24	1365	79.5	1215	9	AF391086	AF391086
25	1347.5	79.1	1105	9	AY042231	AY042231
26	1342	77.6	1083	9	AY042225	AY042225
27	1315.5	77.5	1143	9	AY042230	AY042230
28	1314.5	77.4	1243	9	AY042236	AY042236
29	1313	75.5	939	9	AY042237	AY042237
30	1280.5	71.7	939	9	AY042237	AY042237
31	1215.5	67.9	811	9	AY042237	AY042237
32	1151	67.7	792	9	AY042239	AY042239
33	1148	67.5	1472	9	AY042239	AY042239
34	1144	65.6	807	9	AY042239	AY042239
35	1112.5	63.6	143619	9	AC008812	AC008812
36	1078.5	62.8	4506	9	AF209480S2	AF209480S2
37	1065.5	62.8	190628	2	AC024700	AC024700
38	1054.5	62.2	2391	9	AY042232	AY042232
39	971.5	56.9	994	9	AF209479	AF209479
40	964.5	52.1	4424	9	AC008812	AC008812
41	883.5	51.7	143619	9	AC008812	AC008812
42	877	51.7	17762	9	AF373409	AF373409
43	877	43.5	1220	10	AF373409	AF373409
44	720	37.2	927	10	AF373409	AF373409
45	631	37.2	927	10	AF373409	AF373409

RESULT 1

ALIGNMENTS


```

Db      485 TTGAGACAGCAA 996
RESULT 6
LOCUS   AP166105 1212 bp  DNA  linear  FA: 17-Oct-2001
DEFINITION Sequence 9 from patent US 6280953.
ACCESSION AP166105
VERSION  AP166105.1 GI:16241277
KEYWORDS
SOURCE   Unknown.
ORGANISM Unknown.
REFERENCE
  1 (bases 1 to 1212)
  Messier W. and Sikela J.M.
  Methods to identify polynucleotide and polypeptide sequences which
  may be associated with physiological and medical conditions
  Patent: US 6280953-A 9 28-AUG-2001;
  Location/Qualifiers
    source          1..1212
                   /organism="Unknown"
BASE COUNT  314 a 321 c 348 g 229 t
ORIGIN
Alignment Scores:
Pred. No.      2,050-102      Length:      1212
Score:         1412.50      Matches:      282
Percent Similarity: 77.94%      Conservative: 20
Best Local Similarity: 71.94%      Mismatches:  19
Query Match:   83.28%      Indels:      71
                        Gaps:      4
US-09-831-458A-12 (1-325) x AP166105 (1-1212)
QY      1 MetSerAspSerLysGluPheArgValGlnGlnLeuGlyLeu----- 15
Db      1 ATGAGTACTCCAGAGACCAAGACTGTAGTGGTGGTCTGGAGAGAAATATG 60
QY      16 -----GlycLeuGlyHis 20
Db      61 AGAGGCTTGATTCGGACAGACTGAGGATACAGAGCTTAGCGAGTCTGGCCAT 120
QY      21 G1A1A1eUvAl1eUglnLeuLeuSerTherMetLeuLeuAlaGlyVal1eUvAl1e 40
Db      121 GATGCCCCGAGCTGACACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 166
QY      41 LeuValGlnValSerLysValProSerSerLysSerGlnGlnInsGlnAspAla 60
Db      169 CTGTCCAGAGTGTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 228
QY      61 T1eTyrGlnAsnLeuThrGlnLeuLysAlaAlaValGlyGlnLeuSerGlnLys 80
Db      229 ATCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 288
QY      81 LeuGlnGlnLeuLeuGlnGlnLeuLeuGlnGlnGlnGlnGlnGlnGlnGln 100
Db      289 CTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 348
QY      101 LysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 120
Db      349 AATCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 408
QY      121 LeuTrpGlnLysSerLysLeu----- 127
Db      409 CTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 468
QY      127 ----- 127
Db      469 G1G1G1G1G1G1G1G1G1G1G1G1G1G1G1G1G1G1G1G1G1G1G1G1G1 228
QY      128 -----GlnGlnLysLysGlnGlnLeu 134
Db      529 AAGGCTGTAGTGGTGGAGAGCTTCCAGAGAAATCTAAGCAGCAGAGATCTTAC 588

```

```

QY      135 ThrAlGlnLysAlaAlaValGlyGlnLeuProGlnLysSerLysLeuGlnGln 110
Db      589 AAGGCTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1164
QY      155 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 130
Db      649 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 136
QY      175 GlnLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 150
Db      709 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 156
QY      195 ProLysAspTrpPhePheGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 170
Db      769 GCTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 176
QY      215 ThrLysAspSerVal1eUAlaGlyGlnGlnGlnGlnGlnGlnGlnGlnGln 190
Db      829 TGGTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 196
QY      235 AlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 210
Db      889 GCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 216
QY      255 GlyLeuSerLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 230
Db      949 GAGATTCAATCTAATCTAATCTAATCTAATCTAATCTAATCTAATCTAAT 236
QY      275 ProSerPheGlnArgTyrTrpAsnSerLysGlnGlnGlnGlnGlnGlnGln 250
Db      1009 CCGAGCTTCAAGCAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 256
QY      295 AlaGlnPheSerGlySerGlyTyrPheSerPheSerPheSerPheSerPhe 270
Db      1069 GCGAATTTAGTGGCAATGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 276
QY      315 CysLysLysProAlaAla-----CysPheArgAspGln 325
Db      1129 TGCMAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1164
RESULT 7
AX342003
LOCUS   AX342003 1212 bp  DNA  linear  FA: 17-Oct-2001
DEFINITION Sequence 9 from Patent W0196603.
ACCESSION AX342003
VERSION  AX342003.1 GI:18137981
KEYWORDS
SOURCE   human
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eu-
          Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; H
REFERENCE
  1
  Messier W. and Sikela J.M.
  Methods to identify polynucleotide and polypeptide sequences
  may be associated with physiological and medical condit
  Patent: W0196603-A 9 20-Dec-2001;
  Location/Qualifiers
    source          1..1212
                   /organism="Homo sapiens"
                   /db_xref="taxon:9606"
BASE COUNT  314 a 321 c 348 g 229 t
ORIGIN
Alignment Scores:
Pred. No.      2,050-102      Length:      1212
Score:         1412.50      Matches:      282
Percent Similarity: 77.94%      Conservative: 20
Best Local Similarity: 71.94%      Mismatches:  19
Query Match:   83.28%      Indels:      71
                        Gaps:      4
US-09-831-458A-12 (1-325) x AX342003 (1-1212)

```



```

FEATURES
Source      Location/Qualifiers
            1..1212
            /organism="Pan troglodytes"
            /db_xref="taxon:9598"
            /cell_type="monocyte-derived dendritic cells"
gene        1..>1212
            /gene="DC-SIGN"
            /gene="DC-SIGN"
            /note="HIV-1 trans-receptor, CD209"
            /codon_start=1
            /product="dendritic cell-specific ICAM-3 grabbing
            non-integrin"
            /protein_id="AAK97458.1"
            /translation="MSDSKEPRLOQIGLEBEQLKGLGFQNGYKSLAGLGHPLV
            LQLLSFTLLAGLIVSVKPSYISOESRODYIYONITOLKAAGELSEKSLQELIYQ
            ELTOLKAAGELPEKSKQELIYQELITRLKAAGELPEKSKQELIYQELITRLKAAGEL
            PEKSKQELIYQELITRLKAAGELPEKSKQELIYQELITRLKAAGELPEKSKQELIYQ
            LTRKAAGELPEKSKQELIYQELITOLKAAGELPEKSKQELIYQELITOLKAAGELPEKSKQELIYQ
            WDSITACKRVGQIVIKSAEONLTQSSRSNRPFTWNGSDINEGTQWQWDGSP
            LIPSFQYNNKCEPNNVCEDECAEISGNMNDKCNLAKFWLCKNSAASCSHDEQFL
            SPAPATPPPPA"
BASE COUNT      315 a      316 c      351 g      230 t
ORIGIN
Alignment Scores:
Pred No.      8 72e-102      Length:      1212
Score:         1404.50      Matches:      278
Percent Similarity: 76.79%      Conservative: 23
Best local Similarity: 70.92%      Mismatches: 20
Query Match:    82.81%      Indels:      71
                        Gaps:      4
US-09-831-458a-12 (1-325) x AF391085 (1-1212)
QY      1 MetSerAspSerLysGluProArgValGlnGlnLeuSlyLeuLeu-----15
Db      1 ATGAGTACTGTCAGAGCAAGCAAGCTGACGACGACGAGGCGCTGCTTGGCCAT 60
QY      16 -----GlycylsLeuGlyHis 20
Db      61 AGAGGCTTGATTCGACAGCAATCGAGCTTACAGAGCTTACAGAGCTTCTTGGCCAT 120
QY      21 GYAlaLeuValLeuGlnLeuLeuSerPheMetLeuLeuAlaGlyValLeuValAlaIle 40
Db      121 GCTCCCTCGTGGTGGTCAACACGCTCTCTTACGCTCTTGGCTGGG-----CTG 168
QY      41 LeuValGlnValSerLysValProSerSerLeuSerGlnGlnGlnSerGlnGlnAspAla 60
Db      169 CTGTGCTCAAGTGTCCAGAGCTTCTTAACTTCAATATTAAGTAAAGATATTCAGACAGC 228
QY      61 LLeTyrGlnAsnLeuThrGlnLeuLysAlaAlaValGlyGlnLeuSerGlnLysSerLys 80
Db      229 ATTACAGCAAGCTGACCTTAAAGTTGTATTTGAGTAACTTCTTCAAGAGAAATCTAAG 288
QY      81 LeuGlnGlnLysLeuGlnGlnLeuThrGlnLeuLysAlaAlaValGlyGlnLeuProGln 100
Db      289 CTGCAGAGAGATCTACAGAGAGCTGACCTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAG 348
QY      101 LysSerLysLeuGlnGlnLysLeuThrArgLeuLysAlaAlaValGlyGln 120
Db      349 AATATCAAGCAGCAGATCTACAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 408
QY      121 LeuProGlnLysSerLysLeuGlnGlnLysLeuThrArgLeuLysAlaAlaValGlyGln 140
Db      409 CTTCAGAGAAATCTAGATGTAGAGAGATCTAGAGAGTGTAGCTGAGAGCTGAGAGCTGAG 468
QY      141 ValGlyGlnLeuProGlnLysSerLysLeu-----150
Db      469 GTGGGTGAGCTTCCAGAGAAATCTAAGATCTAGTAACTTACAGAGAGAGTGAAGTGAAG 528
QY      150 -----150

```

```

Db      529 AACCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 548
QY      151 -----151
Db      589 ACCGAGTTGAAGGCTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 648
QY      155 GlnGlnLeuThrArgLeuLysAlaAlaValGlyGlnLeuProGln 668
Db      649 CAAAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 708
QY      175 GlnLysLeuGlnGlnLeuThrAspLeuLysThrAlaPheGlnArgLeuVal 728
Db      709 GAGATCTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 768
QY      195 ProLysAspTrpTrpPhePheGlnGlnLysAspGlySerPheSerAspGly 828
Db      769 CCGTGGAGATGACATCTCTTCAAGAGAACTGTACTTCAATGCTGAGCT 848
QY      215 TrpHisAspSerValThrAlaGlyGlnGlnValArgAlaGlnLeuValVal 868
Db      849 TGGCAGCAATCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 908
QY      235 AlaGlnGlnGlnAsnPheLeuGlnGlnLeuThrThrSerArgSerAsnArgPhe 928
Db      889 GTTAGGAGATAGAACTTCTCAATATTTCAATATTTCAATATTTCAATATTTCAATATTTCA 948
QY      255 GlyLeuSerAspLeuAsnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1008
Db      949 GCACTTCAATCTAATATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1068
QY      275 ProSerPheGlnArgTrpTrpAspSerGlyGlnLeuProAsnAsnSerGlyVal 1088
Db      1009 CCGACCTTCAATATTTCAATATTTCAATATTTCAATATTTCAATATTTCAATATTTCAATAT 1148
QY      295 AlaGlnPheSerGlySerGlyTrpAspAspAsnArgGlySerValAspAla 1168
Db      1069 GCGAGATTTAGTGGCAGATGAGCTGAGATGATGAGCAATCTAATCTTGGAGAG 1228
QY      315 CysLysLysProAlaAla-----CysPheArgAspGln 325
Db      1129 TGCAGAAAGTCCGAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1188

```

Search completed: December 7, 2002, 10:12:41
 Job time : 2737 secs

•
•
•
•

GenCore version 5.1.3
Copyright (c) 1993 2002 Compugen Ltd

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2002, 11:28:11, Search time 2000 Seconds

(without alignments)
2555.113 Million cell updates/sec

Title: US-09-831-458A-12

Percent score: 1696

Sequence: 1 MSDSKPRVVOGLLGLGICR.....RCIVDNYWICKRACRERDE 325

Scoring table:
BLSDUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus_p2n.model DEY=xlh
-O=/cgn2-1/OSPro/spool/US09831458/runat_05122002_103219_17293/FEF_juery.fasta_1_519
-DB=EST -QFMT=fastlap -SUFFIX=std.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -STAR1=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR.SCORE=pcr -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTENT=pio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=0509831458 @CGN 1.1.763 -runat_05122002_103219_17293 -NCPU=6 -FTPT=3
-NO_XLPHY -NO_MMAP -LARGEDUPERR -NEG.SCORES=0 -WAIT -T=NGTlog -DEY_TIMEOUT=120
-WARN_TIMEOUT=10 -THRAPAS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -Delop=6 -DELEXT=7

Database :
1: EST:
2: em_estdb:
3: em_esthm:
4: em_estim:
5: em_estov:
6: em_estpl:
7: em_estro:
8: em_dtc:
9: qb_est1:
10: qb_est2:
11: qb_hic:
12: qb_est3:
13: qb_est4:
14: qb_est5:
15: em_estfun:
16: em_estom:
17: qb_gss:
18: em_gss_hum:
19: em_gss_inv:
20: em_gss_pln:
21: em_gss_vrt:
22: em_gss_fun:
23: em_gss_mam:
24: em_gss_mus:
25: em_gss_other:
26: em_gss_pro:
27: em_gss_rtd:

Pred. No. is the number of results predicted by chance
score greater than or equal to the score of the result
and is derived by analysis of the total score distribution

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Id
1	849	50.1	792	9	AL544658
2	827.5	48.8	828	10	BE562964
3	827	48.8	769	13	B1828875
4	763	45.0	786	14	B1827056
5	739	43.6	706	10	AV445756
6	714	42.1	503	14	R98113
7	711.5	42.0	1136	11	AK007654
8	710	41.9	792	9	AI140165
9	623	39.7	669	10	BM365927
10	559.5	33.0	654	13	BM365285
11	520.5	30.7	441	9	AA914211
12	434.5	25.6	655	10	BM683702
13	428.5	25.3	1165	11	AK007638
14	424.5	25.0	510	14	BM462374
15	424	25.0	833	9	AL570816
16	421.5	24.9	438	9	AL507526
17	417	24.6	1045	11	AK010071
18	407.5	24.0	614	10	BM637496
19	400	23.6	419	14	H90968
20	394	23.2	490	9	AL548311
21	391	23.1	359	10	AV456679
22	384	22.6	1211	11	BP921323
23	379.5	22.4	398	9	AA510960
24	369.5	21.8	284	9	AL614060
25	358.5	21.1	718	10	BE374084
26	343	20.2	586	9	AA543874
27	330.5	19.5	633	10	BM665930
28	330	19.5	395	9	AL308251
29	323	19.0	486	14	BM0140
30	308	18.2	1197	11	BM014779
31	302.5	17.8	429	9	AA165054
32	296	17.5	433	9	AL296429
33	292.5	17.2	434	10	BE361665
34	291.5	17.0	557	12	BM712387
35	289	17.0	816	14	BM018186
36	287	16.9	953	14	BE116245
37	277.5	16.4	597	12	BE512369
38	275	16.2	774	14	B1820664
39	274	16.2	500	9	AA048806
40	273.5	16.1	738	14	BM004812
41	272.5	16.1	566	14	BE182055
42	272.5	16.1	945	14	BM446882
43	268.5	15.8	574	12	BM710604
44	268.5	15.8	736	14	B1104514
45	268	15.8	816	12	BE755884

ALIGNMENTS

RESULT 1
AL544658
LOCUS: AL544658 LTI-NFL006.F12 Homo sapiens cDNA clone
DEFINITION: Prime, mRNA sequence.
ACCESSION: AL544658
VERSION: AL544658.1 GI:12877146
KEYWORDS: EST.
SOURCE: human.
ORGANISM: Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE: Li, W.B., Gruber, C., Josses, J. and Polayes, D. Full-length cDNA libraries and normalization


```

source
1..706
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GLCAP05"
/clone_1lb="GLC"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk (-), Site_1: EcoRI; Site_2: XhoI"

BASE COUNT      168 a      176 g      162 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:      5.86e-77      Length:      706
Score:          739.00      Matches:      128
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      43.57%      Indels:      0
DB:                10      Gaps:      0

US-09-831-458A-12 (1-325) x AV645755 (1-706)
QY 188 GUATGLeuGcysATgHIScysProLysAspTrpIhrPhePheGlnGlyAsnGcysTyrPhe 207
DB 145 GAAGCCCGCTGCGCGGACATGTCGACAGACTGACATCTTCTCCAGGAAACTGTTACTTC 204
QY 208 MetSerAsnSerGlnATgATgATrPHisAspSerValThrAlaGcysGlnGlnValAAGAla 227
DB 205 ATGCTTACTCTCCAGCCGACACTGATAGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 264
QY 228 GlnLeuValAlIleLysThrAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 247
DB 265 CAGCTCGTGGTATGTAAGAACTGCTGAGAGAGAGAACTTCTACAGCTGCGAGATTTCAAG 324
QY 248 SerAsnATgPheSerTrpMetGlyLeuSerAspLeuAsnGlnGlnGlnGlnGlnGlnGlnGln 267
DB 325 AGTACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 384
QY 268 ValAspGlySerProLeuSerProSerPheGlnATgTrpPAsnSerGlyTrpPAsnSerGly 287
DB 385 GTGACGAGCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 444
QY 288 AsnSerGlyAsnGlnAspGlyAlaGlnPheSerGlySerGlyTrpPAsnSerGlyTrpPAsn 307
DB 445 AATAGCGGAAATGAAGACTGTGGGAAATTAGTGGCAGTGGAGTGGAAATGAATGATATG 504
QY 308 AspValAspAsnTrpTrpIleGcys 315
DB 505 GACGTTGACCAATTACTGGAAGCTGG 528

RESULT 6
R98113      503 bp      mRNA      linear      EST 11-SEP-1995
LOCUS      R98113
DEFINITION y966a02.r1 Soares fetal liver spleen INFUS Homo sapiens cDNA clone
IMAGE-200714 5' similar to SP:AA65274 AA6274 HIV CR120-BINDING
C-TYPE ELECTIN - 1, mRNA sequence.
ACCESSION  R98113
VERSION    R98113.1 GI:984774
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
EXTRACTOR  Metacoat; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE  1 (bases 1 to 503)
AUTHORS    Hillier, L., Clark, N., Dubouche, J., Ellisston, K., Hawkins, M., Holman
           M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
           Kitzin, L., Kohling, I., Soares, M., Tan, F., Trevasakis, E., Waterston
           R., Williamson, A., Wohlmann, P., and Wilson, R.
TITLE      The Washu-Merck EST Project
JOURNAL    Unpublished (1995)
COMMENT    Contact: Wilson RK
           Washington University School of Medicine

```

```

4444 Forest Park Parkway, Box 6001, St. Louis, MO
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@washington.wustl.edu
Insert Size: 1845
High quality sequence stops: 470
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through the
IMAGE Consortium (info@image.llnl.gov) for further
insert length: 1845 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 470.

FEATURES
source
1..503
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GLCAP05"
/clone_1lb="GLC"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk (-), Site_1: EcoRI; Site_2: XhoI"

BASE COUNT      126 a      137 c      128 g      105 t      7 others
ORIGIN

Alignment Scores:
Pred. No.:      3.07e-74      Length:      503
Score:          714.00      Matches:      152
Percent Similarity: 95.90%      Conservative: 1
Best Local Similarity: 94.29%      Mismatches: 1
Query Match:      42.10%      Indels:      1
DB:                14      Gaps:      0

US-09-831-458A-12 (1-325) x R98113 (1-503)
QY 188 GUATGLeuGcysATgHIScysProLysAspTrpIhrPhePheGlnGlyAsnGcysTyrPhe 207
DB 145 GAAGCCCGCTGCGCGGACATGTCGACAGACTGACATCTTCTCCAGGAAACTGTTACTTC 204
QY 208 MetSerAsnSerGlnATgATgATrPHisAspSerValThrAlaGcysGlnGlnValAAGAla 227
DB 205 ATGCTTACTCTCCAGCCGACACTGATAGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 264
QY 228 GlnLeuValAlIleLysThrAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 247
DB 265 CAGCTCGTGGTATGTAAGAACTGCTGAGAGAGAGAACTTCTACAGCTGCGAGATTTCAAG 324
QY 248 SerAsnATgPheSerTrpMetGlyLeuSerAspLeuAsnGlnGlnGlnGlnGlnGlnGlnGln 267
DB 325 AGTACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 384
QY 268 ValAspGlySerProLeuSerProSerPheGlnATgTrpPAsnSerGlyTrpPAsnSerGly 287
DB 385 GTGACGAGCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 444
QY 288 AsnSerGlyAsnGlnAspGlyAlaGlnPheSerGlySerGlyTrpPAsnSerGlyTrpPAsn 307
DB 445 AATAGCGGAAATGAAGACTGTGGGAAATTAGTGGCAGTGGAGTGGAAATGAATGATATG 504
QY 308 AspValAspAsnTrpTrpIleGcys 315
DB 505 GACGTTGACCAATTACTGGAAGCTGG 528

RESULT 7

```

	JOURNAL	AF097656	11 kb PCR	mRNA	linear	Apr 19, 2003
	DEFINITION	this musculus 10 day old male pancreas cDNA library full-length enriched library clone: AB000622 b-actin-PKRXA-associated LECTIN TYPE-C (PROBABLY MANNOSYL-BINDING C-TYPE LECTIN DC SIGN), full insert sequence.				
	ACCESSION	AK0007656				
	VERSION	AK0007656.1	GI:12841335			
	KEYWORDS	HTC; CAP trapper.				
	SOURCE	Mus musculus (strain C57BL/6J) 10 day old male pancreas cDNA to mRNA, clone: Jb-RKEN full-length enriched mouse cDNA library clone: AB0067022				
	ORGANISM	Mus musculus				
	REFERENCE	Eukaryotic Metazoa, Chordata, Cladonia, Vertebrata, Euteleostomi, Mammalia; Eutheria; Rodentia; Sciurognathi, Muridae; Murinae, Mus				
	AUTHORS	Carninci, P. and Hayashizaki, Y				
	TITLE	High-efficiency full-length cDNA cloning				
	JOURNAL	Nat Biotechnol 17(12):1044-1048 (1999)				
	MEDLINE	99274254				
	PUBMED	10349636				
	REFERENCE	2				
	AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komio, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y				
	TITLE	Normalization and subtraction of cap-trapper selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
	JOURNAL	Genome Res 10(10):1617-1626 (2000)				
	MEDLINE	20493374				
	PUBMED	11042159				
	REFERENCE	3				
	AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sakaki, N., Carninci, P., Komio, H., Akiyama, J., Nishi, K., Kitazumi, T., Tashtiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishibe, T., Harada, A., Yamamoto, F., Motomizu, H., Sakaguchi, S., Ikuyama, T., Yoshitani, N., Fujiwara, S., Inoue, K., Tojawa, Y., Izawa, M., Ohara, E., Wachihi, M., Yoneda, Y., Ishikawa, T., Oyawa, K., Tanaka, T., Matsuno, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y				
	TITLE	RKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 484 multiplexillary sequencer				
	JOURNAL	Genome Res 10(11):1757-1771 (2000)				
	MEDLINE	20530913				
	PUBMED	11076861				
	REFERENCE	4				
	AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Furukoshi, Y., Komio, H., Adachi, J., Fukuda, S., Alizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamano, K., Saito, T., Okazaki, Y., Toyokuni, T., Bono, H., Kasakawa, T., Sato, R., Kodora, K., Matsuda, H., Ashburner, M., Bateman, S., Casavant, T., Fleischmann, W., Gasteigler, I., Gissi, C., King, R., Kohwi-Mori, H., Knebel, P., Lewis, S., Matsuo, Y., Nikaido, I., Pascoli, V., Quackenbush, J., Schmitt, L.M., Staudt, F., Suzuki, R., Tomita, M., Wagner, T., Washio, T., Sakai, S., Okida, J., Enomoto, M., Arima, H., Baldarelli, R., Barsh, G., Blakes, J., Brillet, D., Boujard, N., Carlinici, P., de Bonaldo, M. F., Brownstein, M. J., Poll, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustintch, S., Hill, D., Holman, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, C., Mazzarelli, S., McDevitt, E., Nielsen, F., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Segal, T., Shibata, Y., Storek, R.F., Suzuki, H., Togo, O.K., K., Wang, K.H., Weitz, C.G., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohzuki, S. and Hayashizaki, Y.				
	TITLE	Functional annotation of a full-length mouse cDNA collection				
	JOURNAL	Nature 409(6821):685-690 (2001)				
	MEDLINE	21085660				
	PUBMED	11217851				
	REFERENCE	5 (bases 1 to 1136)				
	AUTHORS	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,				

TITLE	TITLE					
COMMENT	COMMENT					
FEATURES	FEATURES					
SOURCE	SOURCE					
CDS	CDS					
BASE COUNT	BASE COUNT					
ORIGIN	ORIGIN					
Alignment Scores:						
Pred. No.:	2,386-74					
Score:	711.50					
Percent Similarity:	61.39%					
Best Local Similarity:	46.52%					
Query Match:	41.95%					
DB:	11					
	Gaps: 8					

```

Db      42 TTGGACGATCTCGGACACACAGCAAGCTCCCTGGCTCGAGCTGCTCTCTCTG 151
QY      42 LeuLeuAlaGlyValLeuValAlaIleLeuValGluValSerLeuValPheSerLeu 51
Db      152 TTTTGGCTGGGCTGCTGCTGATCATCTTTTCCAACTGCTCAAAACGGCAAT 205
QY      52 SerGlnGlnGlnSerGlnGlnGlnAlaIleTyrGlnAsnLeuThrGlnLeu 68
Db      206 ACCGAGGAGGAGCAAGCAATCAAGAAATCTCTCAAAATCAAGCTCAATCAAG 265
QY      69 LysAlaAlaValGlyGlnLeuSerGlnLeuSerGlnLeuSerGlnLeuGlnLeu 85
Db      266 ATTACGTCAGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 325
QY      86 GlnGlnLeuThrGlnLeuLeuLysAlaAlaValGlyGlnLeuProGlnLeuSerLysLeuGln 105
Db      326 GACCAACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 355
QY      106 GlnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 125
Db      356 TTTTCCAGATATTTCTTCAATCAATCAATCAATCAATCAATCAATCAATCAAT 394
QY      126 LysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 145
Db      395 ATACCAAGCAAGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 436
QY      146 GlnLysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 165
Db      437 GATCTTCTTCCAAAGCTCC 457
QY      166 GlnLeuPro-----AspGlnSerLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 183
Db      458 AGCTTCCCGGAAAGATCAATCTTCAACAGCAAGCAATCAATCAATCAATCAATCAAT 517
QY      184 TysThrAlaPheGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 203
Db      518 AAGCTCAACTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 577
QY      204 AsnGlySerPheMetSerAsnSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 223
Db      578 AATTGTTACTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 637
QY      224 GluValArgAlaGlnLeuValAlaIleTyrPheAlaGlnGlnGlnGlnGlnGlnGln 243
Db      638 GAAGTGAAGGCTCAAGTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 694
QY      244 GluThrSerArgSerAsnArgPheSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 263
Db      695 CATACTTCAAGCTTAAGCTTAAGCTTAAGCTTAAGCTTAAGCTTAAGCTTAAGCTTA 754
QY      264 ThrTyrGlnTyrValAlaSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 283
Db      755 ACGTGGCTGCTGGGTAATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 814
QY      284 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 303
Db      815 GGGGAGGCTCAAGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 874
QY      304 AspAsnArgLysAspValAlaSerPheTyrPheGlnGlnGlnGlnGlnGlnGlnGln 319
Db      875 GACCTCAATCTGACCTCAAAAGTCTGATCTCAAAAGTCTGCA 922

```

RESULT #
 A0140165
 LOCUS A0140165 PLACE2 Homo sapiens cDNA clone PLACE2000062 5', mRNA
 DEFINITION
 sequence.
 A0140165
 ACCESSION
 A0140165 1 GI-11001686
 VERSION
 KEYWORDS
 EST.
 SOURCE
 ORGANISM
 Human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE
AUTHORS
1 (bases 1 to 792)
Ota, T., Suzuki, Y., Ito, S., Nishikawa, T., Yamamoto, Y., Nakamura, Y.,
Nishikawa, T., Nakamura, Y., Sugano, S., Masuhara, Y.,
HRI human cDNA project (ed. T. Suzuki, Y. Saito,
Yamamoto, Y., Sugiyama, I., Nishikawa, T., Nakamura,
Masuno, Y., Ito, S.)
Unpublished (2000)
CONTACT
Contact: Takao Itoai
Genomics Laboratory
Helix Research Institute
1532-3 Yata, Kisarazu, Chiba 272-0812, Japan
TEL: 81 438-52-3975
FAX: 81 438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project: 5' & 3'-end one pass seq.
Research Institute cDNA library construction:
Virology, Institute of Medical Science, University
of Tokyo
Helix Research Institute
LOCATION/Qualifiers
SOURCE
1. 792
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="PLACE200062"
/clone_id="A0140165"
/tissue_type="placenta"
/note="Vector: pBLSP3"
BASE COUNT 214 a 191 c 244 g 147 t
ORIGIN
Alignment Scores:
pred. No.: 1,966 74 Length: 792
Score: 710.00 Matches: 164
Percent Similarity: 64.47% Conserved: 12
Best Local Similarity: 66.67% Mismatches: 31
Query Match: 41,868 Indels: 66
DB: Gaps: 4
US-09-831-458a-12 (1-325) x A0140165 (1-792)
QY 1 MetSerAspSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1
Db 8 AAGTGAAGGCTCAAGTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1
QY 21 GlnAlaGlnValLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1
Db 52 ----- 1
QY 41 LeuValGlnValSerLysValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1
Db 53 -----GTTGTCAGGCTGGGTAATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1
QY 61 IleTyrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1
Db 104 ATCTATCAAAATCAAGTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1
QY 81 LeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1
Db 164 CTGTAAGGAGATCAAGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1
QY 101 LysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1
Db 224 AATTCAGGCTGGGTAATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1
QY 121 LeuThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1
Db 284 TTGCTAGGAGATCAAGTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1
QY 141 ValGlnGlnLeuProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1
Db 344 GTGCTAGGCTGGGTAATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1
QY 161 LysAlaAlaValGlyGlnLeuSerGlnLeuSerGlnLeuSerGlnLeuGlnLeu 1

```


[illegible][illegible]

is vector free and at least 20bp in length.
 Insert length: 510 Std Error: 0.00
 Plate: BS320044000 Row: B Column: 5
 Seq primer: CGCCAACTCGGAATTAAACG
 High quality sequence stop: 510.
 Location/Qualifiers
 1..510
 /organism="Bos taurus"
 /strain="Angus"
 /db_xref="taxon:9913"
 /clone="BS320044000R5"
 /clone_lib="Subtracted Lewin Cattle Spleen"
 /sex="female"
 /dev_stage="Adult"
 /note="Judson: Spleen: Vector: pBlueScript SK(+/-); Site 1:
 Error: Site 2: Note: This BS cDNA library was produced by
 subtraction of the original non-normalized bovine spleen
 library with 16,800 previously sequenced clones from a
 bovine placenta cDNA library. The original library was
 constructed as described by Hand et al (2000). Genome
 Research 10(9): 1359-1368."

BASE COUNT 125 a 127 c 145 g 112 t
 ORIGIN

Alignment Scores:

Pred. No.: 6,930-40 Length: 510
 Score: 424.50 Matches: 75
 Percent Similarity: 63.41% Conservative: 29
 Best Local Similarity: 45.71% Mismatches: 55
 Query Match: 25,038 Indels: 5
 DB: 13 Gaps: 2

US-09-831-458a-12 (1-325) x BM362374 (1-510)

QY 158 Thirgyleuysala.....Alavalgyluluprcaasgluetystlnu 175
 DB 28 ACCAGGTTCCGAGTCTGACTGAGTTCAGAGACACAGATCGAATCAACCTGGAGAG 87
 QY 176 ILPTGrlngluleuthraspleuysthralaphedluarpleucysarqhlscyspro 195
 DB 88 ATCTTACACGCTGACCGGAGATGATGATGATGATGATGATGATGATGATGATGAT 147
 QY 196 Lysasptphrphrphrphrphrphrphrphrphrphrphrphrphrphrphrphr 215
 DB 148 CAAATTTGATTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 297
 QY 216 Hisasptphrphrphrphrphrphrphrphrphrphrphrphrphrphrphrphr 235
 DB 208 AGATTCGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 267
 QY 236 Gluuluglulnspheleuglnleuglnthserarqserasnarqphesertipmety 255
 DB 268 GAGAGAGAGAAATTCCTGAGACTTGGTATCCCAATAATATAACCCACCTGGATGGC 327
 QY 256 Louseraspleuasoglnchgllylntfeglnl:fvaiaspcl:serproleser:pro 275
 DB 328 CTGATGATACACCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 387
 QY 276 Scptph:gluairqlytrfpaaserclygliafioasnasercly:asnqluaspcysala 295
 DB 388 AGCTTC-----TGGCAAAAAGGAGAGCCCAACACCAAGGAGATGAGAGAGCTGTCG 438
 QY 296 Gluuphserclysercyltrfpaasnaspaasnarqgysaspaalaspasutyrtrpilesys 315
 DB 439 GAATTCGACCAACGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 498
 QY 416 Lyslysprralia 319
 DB 499 GAAAGAGCTCTGG 510

RESULT 15
 AL570816/c 833 bp mRNA Linear EST 16-FEB-2001
 LOCUS AL570816 . 833 bp mRNA Linear EST 16-FEB-2001

DEFINITION AL570816.L01.NL096.L02 Homo sapiens cDNA clone CG001022YH14.3
 Prime, mRNA sequence.
 ACCESSION AL570816
 VERSION AL570816.1 GI:12927495
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 833)
 AUTHORS Li, W.-B., Gruber, C., Jossee, J., and Polyes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope, Centre National de Sequencage
 Bt 701 31066 Evry Evry-A - France
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr

FEATURES

source
 1..833
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CG001022YH14"
 /clone_lib="LTI_NFL006_PL2"
 /tissue_type="Placenta"
 /note="Vector: pCMVSPORT 6, Site 1: NotI, 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pMWScreen 6
 vector. Library was normalized. Library was constructed by
 Life Technologies' Covant - Feng (and wife) Technologies,
 a division of Invitrogen 9800 Medical Center Drive
 Rockville, Maryland 20850, USA Fax: (1) 301 610 8371
 Email: fliang@lifetech.com URL: <http://fulllength.invitrogen.com>"

BASE COUNT 244 a 194 c 219 g 184 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1,816-39 Length: 833
 Score: 424.00 Matches: 82
 Percent Similarity: 93.26% Conservative: 1
 Best Local Similarity: 92.13% Mismatches: 5
 Query Match: 25,008 Indels: 3
 DB: 9 Gaps: 0

US-09-831-458a-12 (1-325) x AL570816 (1-833)

QY 248 Gluashphleuulnleuc:nltrserarqserasnarqphesertipmety:pro 257
 DB 833 AAGAAATTTCTACAGTG-CAGCAATTCAGAAATAC-CGTTCTCTCGAATGAGACTTTC 716
 QY 258 Aspleuasoglnchgllylntfeglnl:fvaiaspcl:serproleser:pro 277
 DB 775 GAGCAAAATCAGCAAGGAGAGCCCAACACCAAGGAGATGAGAGAGCTGTCG 716
 QY 277 Scptph:gluairqlytrfpaaserclygliafioasnasercly:asnqluaspcysala 297
 DB 715 GAAAGAGAGAAATTCCTGAGACTTGGTATCCCAATAATATAACCCACCTGGATGGC 327
 QY 297 Louseraspleuasoglnchgllylntfeglnl:fvaiaspcl:serproleser:pro 275
 DB 655 TATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 596
 QY 317 Spfoalalalacysphearqspcln 325
 DB 595 GCGGATACCTGCTTNAAGAGATGAA 571

Search completed: December 7, 2002, 14:47:13
 Job time: 2070 secs


```

10 991 CTGCAAAAAGCCGCGAGCGCTGCTTCAGAGAGGAA 1024
RESULT 8
ABL41993
ID ABL41993 standard; DNA: 1643 BP.
XX
AC ABL41993;
XX
DE 11-JUN-2002 (first entry)
XX
DE Nucleotide sequence of human polypeptide HP01347.
XX
XX Antibody: antigen; transmembrane domain protein; HP01347; gene: SS.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 25..915
XX FT /*tag= a
XX FT /product= "HP01347"
XX
XX MO200208416-AL.
XX
XX 31-JAN-2002.
XX
XX 24-JUL-2001: 2001WO-JP06371.
XX
XX 24-JUL-2000: 2000JP-0222743.
XX 24-AUG-2000: 2000JP-0254407.
XX
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
XX Kato S, Nagata N, Fujimura N, Kobayashi M, Ito K, Ishizuka Y;
XX WPI: 2002-195877/25.
XX P-PSDB: ABB09715.
XX
XX Antibody preparation by inoculation of an animal with a vector
XX expressing a fusion protein of an antigen on the C-terminal side of a
XX transmembrane domain for use as drugs, diagnostic reagents and
XX laboratory reagents.
XX
XX Example; Page 23-27, 45pp. Japanese.
XX
XX The specification describes a method of antibody preparation. The
XX method comprises inoculating an animal with a vector expressing
XX a fusion protein having an antigen protein fused to the C-terminal
XX side (extracellular) of a transmembrane domain protein (the
XX N-terminal side of which is intracellular), and then isolating and
XX purifying the antibody from the animal. The antibodies can be used
XX as drugs, vaccines, diagnostic reagents and laboratory reagents. The
XX present sequence encodes a polypeptide, designated HP01347, which was
XX used in the course of the invention.
XX
XX Sequence 1643 BP; 388 A; 429 C; 440 G, 386 T; 0 other:
XX
Alignment Scores:
Pred No 4.97e-115 length: 1643
Score: 1407.00 Matches: 286
Percent Similarity: 76.88% Conservat: 0
Best Local Similarity: 76.88% Mismatches: 1
Query Match: 82.96% Indels: 85
DB: 24 gaps: 2
US-09-831-458A-12 (1-325) x ABL41993 (1-1643)
QY 1 MetSerAspSerLysGluProArgValGlnGluLeuLeuLeuLysGluGlnHis 20
DB 25 ATGAGTGCATCCAAAGCAAGGCTCAGCAGCTGGGCTCCGGGGTCTCTGGCCAT 84
QY 21 GAlaLeuValLeuGlnLeuLeuSerPheMetLeuLysValGlnValLeuValAla 40
DB 85 GCGCGCTGCTGTCAGCACTCTCTTATAGTCTTGGATGGATCTTGGTGGCATG 144

```

```

QY 41 LeuValGlnValSerLysValAlaProSerSerLeuSerGlnGlnSerG 100
DB 146 CTGGHCAAGAGAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 144
QY 61 IleTyrGlnAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 100
DB 205 ATCTACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 144
QY 81 LeuGlnGlnIleLeuTyrGlnGlnLeuLeuLeuLeuLeuLeuLeuLeu 100
DB 265 CTGAGAGAGATTTATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 144
QY 101 LysSerLysLeuGlnGlnIleLeuGlnGlnLeuLeuLeuLeuLeuLeu 100
DB 325 AATTCGAAGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 144
QY 121 LeuProGlnLysSerLysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 100
DB 385 TTGCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 144
QY 141 ValGlnGlnLeuProGlnLysSerLysLeuGlnGlnIleLeuTyrGlnGln 100
DB 445 CTGCTTCAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 144
QY 159 -----
DB 505 AAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 144
QY 159 -----
DB 565 AGGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 144
QY 160 -----
DB 625 CAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 144
QY 175 GlnIleTyrGlnGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 100
DB 685 CAAATTCATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 144
QY 195 ProLysAspTyrPhePheGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 100
DB 745 CCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 144
QY 215 TrpHisAspSerValThrAlaCysGlnGlnValAlaAlaGlnLeuVal 100
DB 805 TGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 144
QY 235 AlaGlnGlnGlnAsnFheLeuGlnGlnGlnGlnGlnGlnGlnGlnGln 100
DB 865 CCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 144
QY 255 GlnLeuSerAspLeuAsnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 100
DB 873 -----
QY 275 ProSerPheGlnArgTyrTrpAsnSerGlnGlnProGlnAsnSerLys 100
DB 874 ---CAGCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 144
QY 294 SalGlnLysSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 100
DB 941 TGGGAGATTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 144
QY 314 CysLysLysProAlaAlaLysPheArgAspGln 325
DB 991 CTGCAAAAAGCCGCGAGCGCTGCTTCAGAGAGGAA 1024
RESULT 9
ABA91224
ID ABA91224 standard; CDNA: 1212 BP.
XX

```


AA091225	
ID	AA091225 standard_cDNA, 1212 bp.
XX	
AC	AA091225:
XX	
UT	04-APR-2002 (first entry)
XX	
DE	Gorilla DC-SIGN cDNA coding region.
XX	
PR	See STR: gorilla_gene_sequences_evolution_AIDS_HIV_128
XX	
OS	Gorilla gorilla.
XX	
PN	W1200196603-AZ
XX	
PB	20-nov-2001
XX	
PF	06-JUN-2001 - 2001WC-HS18310
XX	
Pr	04-nov-2000 - 2000NS-05q1415
XX	
PA	(EVOL-) EVOLUTIONARY GENOMICS LLC.
XX	
PI	Messier W, Sikela JM.
XX	
DR	WP1: 2002-130744/17.
XX	
PT	Determining human or non-human primate polynucleotide or polypeptide sequences associated with a physiological trait and have undergone evolutionary changes, for therapeutic use, involves using statistical methods -
XX	
XX	Example 19; Fig 13; 141pp; English.
XX	
CC	The present sequence is that of the coding region of the gorilla DC-SIGN gene. DC-SIGN is expressed on dendritic cells and is known to provide a mechanism for transport of HIV-1 virus to the lymph nodes. HIV-1 binds to the extracellular portion of DC-SIGN and infects the undifferentiated T cells in the lymph nodes via their CD4 proteins. This ultimately leads to "compromiser" of the immune system and to full-blown AIDS. The invention comprises a comparative genomics approach to identify specific gene changes responsible for differences in functions and diseases distinguishing humans from non-humans, particularly primates including gorilla, orangutan and especially chimpanzee. The evolutionary significance of a nucleotide change is determined by the ratio of the non-synonymous substitution rate (KA) to the synonymous rate (KS) of the nucleotide sequence. Polymorphic-like and polymorphic-like sequences corresponding to evolved traits may be relevant to human diseases or conditions such as unique or enhanced human brain functions, longer human life spans, susceptibility or resistance to disease, including AIDS and cancer, and aesthetic traits such as hair growth. KA/KS ratios for chimpanzee/human, human/gorilla and chimpanzee/gorilla DC-SIGN cDNA sequences are 1.3, 0.87 and 1.4, respectively. It is theorized that chimpanzee resistance to progression to full blown AIDS may be due in part to inability of HIV-1 to bind to chimpanzee DC-SIGN for transport to the lymph nodes. After determining the 3-dimensional structure of DC-SIGN, a rational drug design approach can be used to mimic the effects of chimpanzee DC-SIGN without interfering with the normal functions of human DC-SIGN.
CC	
XX	
SE	Sequence 1212 BP; 315 A; 319 G; 349 C; 229 T; 0 other;
XX	
Alignment Scores:	
Prod. No.:	5,366-114
Scores:	1393.50
Percent Similarity:	76.28%
Best local Similarity:	70.92%
Query Match:	82.16%
DB:	24
	Gaps: 4
	Length: 1212
	Matches: 278
	Mismatches: 21
	Indels: 22
	Gaps: 4

Alignment Scores:

Pred. No.:	8,380-54	Length:	792
Score:	710.00	Matches:	164
Percent Similarity:	64.47%	Conservative:	12
Best Local Similarity:	60.07%	Mismatches:	31
Query Match:	41.86%	Indels:	66
DB:	22	Gaps:	4

US-09-831-458a-12 (1-325) x AAH08239 (1-792)

```

QY 1 MetSerAspSerLysGluProArgValGlnGlnLeuGlyLeuLeuGlyGlySerGlyHis 20
   |||||
DB 8 ATGGTACTCTCCAGGAGACCAAGGTTCCAGCACTGGGCTCTCTG----- 52
QY 21 GlyAlaLeuValLeuGlnLeuLeuSerPheMetLeuLeuAlaGlyAlaValAlaIle 40
   -----
DB 52 ----- 52
QY 41 LeuValGlnValSerLysValProSerSerLeuSerGlnGlnGlnSerGlnGlnAspAla 60
   |||||
DB 53 -----GTGTCAGAGGTCCTCCAGCTCCCTAAGTCACCAACAAATTCAGCAGAGCGCA 103
QY 61 LLeTyrGlnAsnLeuThrGlnLeuLysAlaAlaValGlyLeuLeuSerGlnLysSerLys 80
   |||||
DB 104 ATCTAGCAGAGAACTGACCCAGCACTTAAAGCTACAGCTGGGTGAGCTCTCTCAGAGAAATCCAG 163
QY 81 LeuGlnGlnLLeTyrGlnGlnLeuThrGlnLeuLysAlaAlaValGlnGlnLeuProGln 100
   |||||
DB 164 CTGCAAGCAATCTACCAAGTAACTAACTAACTAACTAACTAACTAACTAACTAACTAACTAA 243
QY 101 LysSerLysLeuGlnGlnGlnLLeTyrGlnGlnLeuThrArgLeuLysAlaAlaValGln 120
   |||||
DB 224 AAATCCAGAGCTGACAGGAGATCTACCAAGAGTAACTAACTAACTAACTAACTAACTAACTAA 283
QY 121 LeuProGlnLysSerLysLeuGlnGlnLLeTyrGlnGlnLeuThrArgLeuLysAlaAla 140
   |||||
DB 284 TTGCGCAGAGAAATCCAGCTACAGAGATCTACCAAGAGTAACTAACTAACTAACTAACTAACTAA 343
QY 141 ValGlnGlnLeuProGlnLysSerLysLeuGlnGlnLLeTyrGlnGlnLeuThrArgLeu 160
   |||||
DB 344 GTGGGTACTGTCAGAGAAATCCAGGAGTAACTAACTAACTAACTAACTAACTAACTAACTAACTAA 403
QY 161 LysAlaAlaValGlnLysLeuProAspGlnSerLysGlnGlnGlnLLeTyrGlnGlnLeu 180
   |||||
DB 404 AAGGCTGAGAGGAGTGTAGTTCAGAGAAATCCAGAGTGTAGTTCAGAGAGAGTGTAGTTCAGAGAG 463
QY 181 ThrAspLeuLysThrAlaPheGlnValLeuLeuGlyArgHisLysSerProLysAspTrpPhe 200
   |||||
DB 464 ACGGAGCTGAGAGGCTGAGAGTGTAGTTCAGAGTGTAGTTCAGAGTGTAGTTCAGAGTGTAGTTC 493
QY 201 PheGlnGlnLysAsnGlyTrpPheMetSerAsnSerGlnArgAsnTrpHisAspSerValThr 220
   -----
DB 494 -----CCAGAGAAATCC 505
QY 221 AlaCysGlnGlnValAlaArgAlaGlnLeuValValIleLysThrAla----- 235
   |||||
DB 506 AAGTCTGAGAGAGATCTACCAAGTACAGGAGTGTAGTTCAGAGTGTAGTTCAGAGTGTAGTTCAG 565
QY 236 -----GlnGlnGlnAsnPheLeuGlnLeuGlnLLeTyrSerArgSerAsnArgPheSer 252
   |||||
DB 566 GACGAGCTGAGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 625
QY 253 TrpMetGlyLeuSerAspLeuAsnGlnGlnGlyThrTrp 265
   |||||
DB 626 -----GAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 652

```

RESULT 14
AAD03221
ID AAD03221 standard; cDNA: 1516 bp.
XX
XX AAD03221:
AC
XX
DT 13-JUN-2001 (first entry)

XX
DE Mouse type II integral membrane protein, CIRE cDNA.
XX
KW Mouse, type II integral membrane protein, CIRE; anti-
KW Immunosuppressive; autoimmune disorder; allergy; viral
KW antigen presenting cell; bacterial infection; macrophage
KW DC; ss.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS /*tag= a
FT PT /product "Mouse type II integral memb"
XX
XX W0200119869-A1.
XX
XX 22-MAR-2001.
XX
XX 13-SEP-2000; 2000WO-A01096.
XX
XX 13-SEP-1999; 99AU-0002788.
XX
XX (CODN-) COUNCIL QUEENSLAND INST MEDICAL RES.
XX
XX Caminschi I, Wright MD, Shortman KD;
XX
XX WPI: 2001-244788/25.
XX
XX F-PSDB; AAV72948.
XX
XX Novel type II integral membrane protein, CIRE, precursor
XX In dendritic cells, macrophages and their precursors se-
XX cret cells, useful for identifying compounds having immuno-
XX activity
XX
XX Claim 6; Fig 1; 53pp; English.
XX
XX
XX The present cDNA sequence encodes mouse type II integra-
XX protein (designated as CIRE) which is preferentially ex-
XX pressed in dendritic cells, macrophages and their precursors. CIRE
XX CIRE protein has a N-terminal cytoplasmic domain, transmem-
XX and a C-type lectin domain in the extracellular region.
XX for screening immunoregulatory compounds and as a marker
XX cell (DC) population. CIRE specific ligand such as mono-
XX are useful for isolating an antigen presenting cell (APC)
XX biological sample. CIRE DNA is also used as tools to ana-
XX properties and functions of the CIRE gene/protein. For ex-
XX molecules are used to regulate animal models, which lack
XX genes and to isolate regulatory regions of the CIRE gene.
XX regions are used to selectively express exogenous genes
XX lipids are useful for modulating immune responses by in-
XX the function, migration or maturation of DC or APC. The
XX and interfere with the function of myeloid dendritic cell
XX antigen processing is undertaken by lymphoid DC which loss
XX suppression and anergy for treating allergies and autoim-
XX The ligands are also useful to block or enhance the inter-
XX viruses or bacteria with DC, for improving resistance to
XX These ligands are also used to target molecules such as
XX components, to DC or APC.
XX
XX Sequence 1516 bp; 430 A; 44 C; 815 G; 426 T; 1 other;
XX
XX
XX Alignment Scores:
Pred. No.: 5,21e-09 Length: 1516
Score: 546.50 Matches: 121
Percent Similarity: 42.82% Conservative: 24
Best Local Similarity: 34.77% Mismatches: 50
Query Match: 32.22% Indels: 149
DB: 22 Gaps: 6
XX
XX US-09-831-458a-12 (1-325) x AAH03221 (1-1516)
QY 1 MetSerAspSerLysGluProArgValGlnGlnLeu-----

[illegible]

XX	AAH99365 standard; cDNA; 433 bp.
AC	AAH99365:
DT	16-OCT-2001 (first entry)
DE	Human protein encoding cDNA sequence SEQ ID NO:200.
XX	
KW	Human: cancer; ulcer; HIV infection; human immunodeficiency virus;
KW	anti-tumour activity; antitubercular; anti-infective; immunosuppressive;
KW	antibacterial; endocrine; cardiac; central nervous system; vitreous;
KW	anti-HIV; fungicide; antimetabolic; cardiovascular; antidiabetic; rheumatoid;
KW	antiparasitic; haemostatic; pulmonary; antiviral; osteoporosis; eczema;
KW	dermatological; antiallergic; antistimulant; antibacterial; cytotoxic;
KW	immunorelaxant; antidepressant; motility; antiparkinsonian; infectious;
KW	immunosuppressant; gene therapy; antisense; therapy; vaccine; inflammation;
KW	antiangiogenic; rheumatoid arthritis; septic shock; pancreatitis;
KW	cardiac dysfunction; neurodegenerative; cardiac anaphylaxis; autoimmune;
KW	genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW	rheumatoid arthritis; osteoporosis; severe combined immunodeficiency;
KW	allergic rhinitis; diabetes; multiple sclerosis; depression;
KW	Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW	neurological disorder; ss.
OS	Homo sapiens.
XX	
PN	WO200153455-A2.
PD	26-JUL-2001.
XX	
PP	22-IPC-2000; 2000WO-US85017.
XX	
PR	23-DEC-1999; 99US-0471275.
FR	21-JAN-2000; 2000DS-0482725.
PR	25-JAN-2000; 2000DS-0552317.
XX	
BA	(HYSE-) HYSPQ INC.
PI	Tang YT, Liu C, Drmanac RJ;
DR	WP1: 2001-457603/49.
NP	P-INSDR; AAM25424.
XX	
FT	Isolated human polynucleotides encoding polypeptides, useful for the treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
XX	
PS	Claim 1; Page 387; 1217pp; English.
XX	
CC	AAP09161 to AAP09904 encode the human Proteins given in AAM25425 to
CC	AAM25466, the proteins can have activities based on the tissues and
CC	cells they are expressed in, such as; antiinflammatory, antirheumatic;
CC	analgesic, immunosuppressive, antibacterial, endocrine, cardiac,
CC	central nervous system, viricide, anti-HIV, fungicide, antimutagen,
CC	cardiovascular; antidiabetic; antiangiogenic; haemostatic; pulmonary;
CC	antiviral; osteoporotic; dermatologically; antiallergic; antasthmatic;
CC	antidiabetic; cytostatic; neuroproliferative; antidepressant; motility;
CC	antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC	encoding them can be used in gene therapy, antisense therapy and vaccine
CC	production. The proteins and polynucleotides are useful for screening for
CC	agonists or antagonists of a protein and for the treatment and diagnosis
CC	of disorders associated with the activity of a protein e.g. inflammation
CC	rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC	neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC	infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC	dementia, pluralist disorders, thrombocytopenia, wounds, burns, ulcers,
CC	dermatological, severe combined immunodeficiency, eczema, allergic
CC	rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC	Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC	neurological disorders.
XX	
SO	Sequence 433 BP; 110 A; 114 C; 128 G; 81 T; 0 other;

•
•
•
•

GenCore version 5.1.4
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: December 7, 2002, 09:41:20 ; Search time per sequence
(without alignments)
787.827 Million cell updates/sec

Title: US-09-831-458a-12

Perfect score: 1696

Sequence: 1 MSDSKPRVQOGLIETGLCH

Scoring table:

RUSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: SPTRMBL_21:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvrvirus:*
16: sp_bacteriophage:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1646.5	97.1	376	4	Q9BXS3
2	1639	96.6	359	4	Q9BXS3
3	1628	96.0	359	4	Q9BXS3
4	1494.5	88.1	332	4	Q9BXS3
5	1412.5	83.3	404	4	Q9BXS3
6	1404.5	82.8	404	4	Q9BXS3
7	1375	81.1	381	6	Q9BXS3
8	1375	81.1	381	6	Q9BXS3
9	1371	80.8	404	4	Q9BXS3
10	1369.5	80.7	404	4	Q9BXS3
11	1365	79.5	381	6	Q9BXS3
12	1347.5	79.5	404	6	Q9BXS3
13	1315.5	77.6	360	4	Q9BXS3
14	1314.5	77.6	360	4	Q9BXS3
15	1303	76.8	334	4	Q9BXS3
16	1274	75.1	380	4	Q9BXS3

17	1215.5	71.7	312	4	Q9BXS3
18	1157	68.2	296	4	Q9BXS3
19	1148	67.7	263	4	Q9BXS3
20	1112.5	65.6	268	4	Q9BXS3
21	901	53.1	232	4	Q9BXS3
22	874	51.5	297	4	Q9BXS3
23	848	50.0	272	4	Q9BXS3
24	720	42.5	335	1	Q9BXS3
25	711.5	42.0	251	4	Q9BXS3
26	711.5	42.0	311	1	Q9BXS3
27	710	41.9	216	4	Q9BXS3
28	631	32.2	235	1	Q9BXS3
29	612	35.1	208	1	Q9BXS3
30	588.5	34.7	237	1	Q9BXS3
31	562	33.1	207	1	Q9BXS3
32	546.5	32.2	238	1	Q9BXS3
33	538.5	31.8	178	1	Q9BXS3
34	503.5	29.7	148	1	Q9BXS3
35	486.5	28.7	211	1	Q9BXS3
36	464.5	27.4	168	4	Q9BXS3
37	428.5	25.3	274	1	Q9BXS3
38	413	24.4	194	1	Q9BXS3
39	369	21.8	331	1	Q9BXS3
40	351	20.7	309	1	Q9BXS3
41	344.5	20.3	309	1	Q9BXS3
42	334.5	19.7	309	1	Q9BXS3
43	313.5	18.5	293	1	Q9BXS3
44	306	18.0	284	1	Q9BXS3
45	303.5	17.9	292	4	Q9BXS3

ALIGNMENTS

Q9BXS3	PRELIMINARY:	PR1:	376 AA.
17	Q9BXS3	Q9BXS3	
18	Q9BXS3	Q9BXS3	
19	Q9BXS3	Q9BXS3	
20	Q9BXS3	Q9BXS3	
21	Q9BXS3	Q9BXS3	
22	Q9BXS3	Q9BXS3	
23	Q9BXS3	Q9BXS3	
24	Q9BXS3	Q9BXS3	
25	Q9BXS3	Q9BXS3	
26	Q9BXS3	Q9BXS3	
27	Q9BXS3	Q9BXS3	
28	Q9BXS3	Q9BXS3	
29	Q9BXS3	Q9BXS3	
30	Q9BXS3	Q9BXS3	
31	Q9BXS3	Q9BXS3	
32	Q9BXS3	Q9BXS3	
33	Q9BXS3	Q9BXS3	
34	Q9BXS3	Q9BXS3	
35	Q9BXS3	Q9BXS3	
36	Q9BXS3	Q9BXS3	
37	Q9BXS3	Q9BXS3	
38	Q9BXS3	Q9BXS3	
39	Q9BXS3	Q9BXS3	
40	Q9BXS3	Q9BXS3	
41	Q9BXS3	Q9BXS3	
42	Q9BXS3	Q9BXS3	
43	Q9BXS3	Q9BXS3	
44	Q9BXS3	Q9BXS3	
45	Q9BXS3	Q9BXS3	

Query Match 97.1% Score 1646.5, 1646.5, 1646.5
Best Local Similarity 86.2% Pred. No. 5.7e-100
Matches 324, Conservative 1, Mismatches 0, 169

```

Db 1 MSDSKREPVQQLGLEEDPTISGTRLEPRPDPOQOIHGKSSGGLGALVLIQILSPML 60
QY 43 IAVIVAAIIYQVSKVPSLSLQFQSDQDAIYQNTQLKAAVGLSKRSKIQTLYQELTQLK 92
Db 61 IAVIVAAIIYQVSKVPSLSLQFQSDQDAIYQNTQLKAAVGLSKRSKIQTLYQELTQLK 120
QY 93 AAVGELPEKSKLOEITQELTRKAAVGLPEKSKLOEITQELTRKAAVGLPEKSKLOE 136
Db 121 AAVGELPEKSKLOEITQELTRKAAVGLPEKSKLOEITQELTRKAAVGLPEKSKLOE 180
QY 147 -----LKAAGVLEPKSKLOEITQELTRKAAVGLPDQSKQOQIYQELTDLKTAEE 189
Db 181 IYQELTRKAAVGLPEKSKLOEITQELTRKAAVGLPDQSKQOQIYQELTDLKTAEE 240
QY 190 LCHRCPEQWTFEFGNCFYFMSNSQPNMHSVTACQEVPAQIVITAEFQNFGLQTSKSN 249
Db 241 LCHRCPEQWTFEFGNCFYFMSNSQPNMHSVTACQEVPAQIVITAEFQNFGLQTSKSN 300
QY 250 RFSMGLSLDNGEGTQWQWVGSPLSPSPFYWNSQEPNNSGNEDEAFESGSGMNRCQV 309
Db 301 RFSMGLSLDNGEGTQWQWVGSPLSPSPFYWNSQEPNNSGNEDEAFESGSGMNRCQV 360
QY 310 DNYWICKKPAACFRDE 325
Db 361 DNYWICKKPAACFRDE 376

```

RESULT 2

```

QYH2X3 ID QYH2X3 PRELIMINARY: PRT: 399 AA.
AC QYH2X3
DT 01-MAR-2001 (TEMBLrel 16, created)
DT 01-MAR-2001 (TEMBLrel 16, last sequence update)
DT 01-JUN-2002 (TEMBLrel 21, last annotation update)
DE Probable mannose binding C-type lectin DC-SIGN (MUC-SIGN2 type I isoform).
GN ABC OR CD209L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
KN 11
KF SEQUENCE FROM N.A.
RX MEDLINE=20472367; PubMed 10975799;
RA Solleux E.J., Barton R., Trowsdale J.;
RT "DC-SIGN, a related gene, DC-SIGNR, and CD23 form a cluster on 19p13."
RT 19p13."
RT J. Immunol. 165:2937-2947(2000).
KN 12
KF SEQUENCE FROM N.A.
RA Solleux E.J.;
RT Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases
KN 13
KF SEQUENCE FROM N.A.
RX PubMed=11337487;
KA Mummidi S., Calano G., Lam L., Hoefle A., Telles V., Bequim K.,
KA Jimenez F., Ahuja S.S., Ahuja S.K.;
RT Extensive repertoire of Membrane-bound and soluble Dendritic Cell-
RT specific ICM-3-grabbing Neutrophin 1 (DC-SIGN) and DC-SIGN2
RT isoforms: Inter-individual Variation in Expression of DC-SIGN
RT Transcripts."
RT J Biol Chem. 276:33196-33212(2001)
DR EMBL: AF209481; AAC13815.2;
DR EMBL: AF209480; AAC13815.2; JOINED.
DR EMBL: AF042344; AAK91859.1;
DR HSSP: P22897; IEGG.
DR InterPro: IPR002353; AntiFceR2E11.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C_1
DR PRINTS: PR00356; ANTI-FCE2E11.
DR SMART: SM00034; CLECT_1
DR PROSITE: PS00615; C-TYPE_LECTIN_1; DNKNWN_1
DR PROSITE: PS00615; C-TYPE_LECTIN_2;

```

```

KW Lectin.
SO SOURCE 499 AA; 45450 MW; 61AIC99722AF594 CIRC64;
Query Match 96.6%; Score 1649; DB 4; Length 399;
Post local similarity 81.5%; Pred. No. 1,9e+99;
Matches 325; Conservative 0; Mismatches 0; Indels 74; Gaps 2;
QY 1 MSDSKREPVQQLGLEEDPTISGTRLEPRPDPOQOIHGKSSGGLGALVLIQILSPML 32
Db 1 MSDSKREPVQQLGLEEDPTISGTRLEPRPDPOQOIHGKSSGGLGALVLIQILSPML 60
QY 43 IAVIVAAIIYQVSKVPSLSLQFQSDQDAIYQNTQLKAAVGLSKRSKIQTLYQELTQLK 92
Db 61 IAVIVAAIIYQVSKVPSLSLQFQSDQDAIYQNTQLKAAVGLSKRSKIQTLYQELTQLK 120
QY 93 AAVGELPEKSKLOEITQELTRKAAVGLPEKSKLOEITQELTRKAAVGLPEKSKLOE 152
Db 121 AAVGELPEKSKLOEITQELTRKAAVGLPEKSKLOEITQELTRKAAVGLPEKSKLOE 180
QY 153 IYQELTRKAAVGLPEKSKLOEITQELTRKAAVGLPDQSKQOQIYQELTDLKTAEE 166
Db 181 IYQELTRKAAVGLPEKSKLOEITQELTRKAAVGLPDQSKQOQIYQELTDLKTAEE 240
QY 167 LPDQSKQOQIYQELTDLKTAEEPLPRPDQWTFEFGNCFYFMSNSQPNMHSVTACQEV 226
Db 241 LPDQSKQOQIYQELTDLKTAEEPLPRPDQWTFEFGNCFYFMSNSQPNMHSVTACQEV 300
QY 227 AAVIVKIAHGHQFQIQLQISKSKFQWMLSLQDQDCTQWQWVGSPLSPSPFYWNSQEP 286
Db 403 AAVIVKIAHGHQFQIQLQISKSKFQWMLSLQDQDCTQWQWVGSPLSPSPFYWNSQEP 460
QY 287 NSNSNLCALFSSGSGMNRCQVQWVGSPLSPSPFYWNSQEPNNSGNEDEAFESGSGMN 325
Db 461 NSNSNLCALFSSGSGMNRCQVQWVGSPLSPSPFYWNSQEPNNSGNEDEAFESGSGMN 399

```

RESULT 3

```

QYH209 ID QYH209 PRELIMINARY: PRT: 399 AA.
AC QYH209
DT 01-MAR-2001 (TEMBLrel 16, created)
DT 01-MAR-2001 (TEMBLrel 16, last sequence update)
DT 01-JUN-2002 (TEMBLrel 21, last annotation update)
DE Probable mannose binding C-type lectin DC-SIGNR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
KN 11
KF SEQUENCE FROM N.A.
RX MEDLINE=20472367; PubMed 10975799;
RA Solleux E.J., Barton R., Trowsdale J.;
RT "DC-SIGN, a related gene, DC-SIGNR, and CD23 form a cluster on 19p13."
RT 19p13."
RT J. Immunol. 165:2937-2947(2000).
KN 12
KF SEQUENCE FROM N.A.
RA Solleux E.J.;
RT Submitted (JAN 2001) to the EMBL/GenBank/DBJ databases.
KN 13
KF SEQUENCE FROM N.A.
RX PubMed=11337487;
KA Mummidi S., Calano G., Lam L., Hoefle A., Telles V., Bequim K.,
KA Jimenez F., Ahuja S.S., Ahuja S.K.;
RT Extensive repertoire of Membrane-bound and soluble Dendritic Cell-
RT specific ICM-3-grabbing Neutrophin 1 (DC-SIGN) and DC-SIGN2
RT isoforms: Inter-individual Variation in Expression of DC-SIGN
RT Transcripts."
RT J Biol Chem. 276:33196-33212(2001)
DR EMBL: AF209481; AAC13815.2;
DR EMBL: AF209480; AAC13815.2; JOINED.
DR EMBL: AF042344; AAK91859.1;
DR HSSP: P22897; IEGG.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C_1
DR ELMETS: FE00356; ANTI-FCE2E11.
DR SMART: SM00034; CLECT_1
DR PROSITE: PS00615; C-TYPE_LECTIN_1;
DR PROSITE: PS00615; C-TYPE_LECTIN_2;
KN 14
KF SEQUENCE FROM N.A.
RX PubMed=11337487;
KA Mummidi S., Calano G., Lam L., Hoefle A., Telles V., Bequim K.,
KA Jimenez F., Ahuja S.S., Ahuja S.K.;
RT Extensive repertoire of Membrane-bound and soluble Dendritic Cell-
RT specific ICM-3-grabbing Neutrophin 1 (DC-SIGN) and DC-SIGN2
RT isoforms: Inter-individual Variation in Expression of DC-SIGN
RT Transcripts."
RT J Biol Chem. 276:33196-33212(2001)
DR EMBL: AF209481; AAC13815.2;
DR EMBL: AF209480; AAC13815.2; JOINED.
DR EMBL: AF042344; AAK91859.1;
DR HSSP: P22897; IEGG.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C_1
DR ELMETS: FE00356; ANTI-FCE2E11.
DR SMART: SM00034; CLECT_1
DR PROSITE: PS00615; C-TYPE_LECTIN_1;
DR PROSITE: PS00615; C-TYPE_LECTIN_2;

```

Query Match 96.6%; Score 1628; DB 4; Length 399;
Post local similarity 81.2%; Pred. No. 9.8e+99;
Matches 325; Conservative 0; Mismatches 1; Indels 74; Gaps 2;

```

UY 1 MSDSKEPPVVOGLGLT-----GCTSHATVAVLSEFME 32
DB 1 MSDSKEPPVVOGLGLEHPPTTSGIRLPPDPHQDIGHKSSSTGCTGHTAVIGITSPML 60
UY 33 LAGVAVAILVSVSPSSLSQSESDAIVQNTOLKAAGVLESEKSLQEIYQELTQCK 92
DB 61 LAGVAVAILVSVSPSSLSQSESDAIVQNTOLKAAGVLESEKSLQEIYQELTQCK 120
UY 93 AAVGELPEKSKLOETIYQELTPIKAAGVLEPEKSKLOETIYQELTPIKAAGVLEPEKSKLOE 152
DB 121 AAVGELPEKSKLOETIYQELTPIKAAGVLEPEKSKLOETIYQELTPIKAAGVLEPEKSKLOE 180
UY 153 IYQELTPIKAAGVLEPEKSKLOETIYQELTPIKAAGVLEPEKSKLOE 166
DB 181 IYQELTPIKAAGVLEPEKSKLOETIYQELTPIKAAGVLEPEKSKLOE 240
UY 167 LPDQSKQOUIYQELTDKTAERKLCRHPKQWIFQGNQVFMNSQNMHDSVIAVQEVK 426
DB 241 LPDQSKQOUIYQELTDKTAERKLCRHPKQWIFQGNQVFMNSQNMHDSVIAVQEVK 300
UY 227 AGVAVIKTAFQNTLQCTSSSNPFSMWGLSDLNQPTWQWVNSPISPSQPYWNSGEP 286
DB 301 AGVAVIKTAFQNTLQCTSSSNPFSMWGLSDLNQPTWQWVNSPISPSQPYWNSGEP 360
UY 287 NNSGNEPCAEFSGSGSWNNPCTDVNNTWICKKPAACFIDE 325
DB 361 NNSGNEPCAEFSGSGSWNNPCTDVNNTWICKKPAACFIDE 399

```

RESULT 4

```

ID 0960P5 PRELIMINARY PRI: 332 AA.
AC 01-DEC-2001 (TREMblrel, 19, Created)
DT 01-DEC-2001 (TREMblrel, 19, Last sequence update)
DT 01-MAR-2002 (TREMblrel, 20, Last annotation update)
DE SDC-SIGN2 type I isoform.
CN CD209L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catartini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11337487;
RA Mummidi S., Catano G., Lam L., Hoefle A., Telles V., Begum K.,
RA Jimenez F., Ahuja S.S., Ahuja S.K.;
RA "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
RT Specific ICAM-3-grabbing Nominogrin 1 (DC-SIGN1) and DC-SIGN2
RT Isoforms: Inter-individual Variation in Expression of DC-SIGN
RT Transcripts.";
RL J. Biol. Chem. 276:33196-33212(2001).
DB EMBL: AY042238; AAK91863.1;
DB InterPro: IPR001304; Lectin_C1;
DB Pfam: PF00059; Lectin_C1;
DB PROSITE: PS00615; G-TYPE-LECTIN_1; UNKNOWN_1;
DB PROSITE: PS00615; G-TYPE-LECTIN_2; 1;
DB SEQUENCE 332 AA; 37953 MW; C6PDEP9471H07406 GR674;

```

```

Query Match 88.1% Score 1494.5. R4 4; Length 332;
Best Local Similarity 85.2% Pred. No. 41e-90;
Matches 293; Conservative 5; Mismatches 15; Gaps 3; Gaps 2.
UY 1 MSDSKEPPVVOGLGLT-----GCTSHATVAVLSEFME 41
DB 1 MSDSKEPPVVOGLGLEHPPTTSGIRLPPDPHQDIGHKSSSTGCTGHTAVIGITSPML 48
UY 42 VQSVSPSSLSQSESDAIVQNTOLKAAGVLESEKSLQEIYQELTQCK 101
DB 49 VQSVSPSSLSQSESDAIVQNTOLKAAGVLESEKSLQEIYQELTQCK 108
UY 102 SKLQETIYQELTPIKAAGVLEPEKSKLOETIYQELTPIKAAGVLEPEKSKLOE 161

```

```

DB 109 SKLQETIYQELTPIKAAGVLEPEKSKLOETIYQELTPIKAAGVLEPEKSKLOE 161
UY 162 AAVGELPEKSKLOETIYQELTPIKAAGVLEPEKSKLOETIYQELTPIKAAGVLEPEKSKLOE 192
DB 169 AAVGELPEKSKLOETIYQELTPIKAAGVLEPEKSKLOETIYQELTPIKAAGVLEPEKSKLOE 228
UY 222 CAEVAQGVVITKTAEPQNTLQCTSSSNPFSMWGLSDLNQPTWQWVNSPISPSQPYWNSGEP 286
DB 229 CAEVAQGVVITKTAEPQNTLQCTSSSNPFSMWGLSDLNQPTWQWVNSPISPSQPYWNSGEP 360
UY 252 NSGEPNNSGNEPCAEFSGSGSWNNPCTDVNNTWICKKPAACFIDE 325
DB 289 NSGEPNNSGNEPCAEFSGSGSWNNPCTDVNNTWICKKPAACFIDE 399

```

RESULT 5

```

ID 09NNX6 PRELIMINARY PRI: 404 AA.
AC 01-OCT-2000 (TREMblrel, 15, Created)
DT 01-OCT-2000 (TREMblrel, 15, Last sequence update)
DT 01-JUN-2002 (TREMblrel, 21, Last annotation update)
DE Membrane-associated lectin type-C (probably mouse hinc-
DE lectin DC-SIGN) (DC-SIGN1 type I isoform).
CN CD209.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catartini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=92390446; PubMed=1518859;
RA Curtis B.M., Scharnowski S., Watson A.J.;
RA "Sequence and expression of a membrane-associated C-type
RT exhibits CD4-independent binding of human immunodeficient
RT envelope glycoprotein gp 120."
RL Proc Natl Acad Sci U S A 89:4456-4460(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX PubMed=2432267; PubMed=10975799;
RA Solleux E.J., Barten R., Trowsdale J.;
RA "DC-SIGN, a related gene, DC SIGNR, and CD24 form a cluster
RT 19p13."
RL J. Immunol. 165:2937-2942(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX PubMed=21157496; PubMed=1125714;
RA Bashirova A.A., Geltenbock B.H., van Duinhoven G.C.J.,
RA van Vliet S.T., Flierberg R.G., Martin M.P., Wu L., Mar-
RA Vleeling N., Koelle P.A., Kwak-Kwani V.N., van Kooyk Y.;
RA "A dendritic cell specific intercellular adhesion molecule-
RT nonintegrin (DC-SIGN)-related protein is highly expressed
RT in liver sinusoidal endothelial cells and promotes HIV-1 in-
RT J. Exp. Med. 193:671-678(2001).
RN [4]
RP SEQUENCE FROM N.A.
RX PubMed=11337487;
RA Mummidi S., Catano G., Lam L., Hoefle A., Telles V., Begum K.,
RA Jimenez F., Ahuja S.S., Ahuja S.K.;
RA "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
RT Specific ICAM-3-grabbing Nominogrin 1 (DC-SIGN1) and DC-SIGN2
RT Isoforms: Inter-individual Variation in Expression of DC-
RT Transcripts.";
RL J. Biol. Chem. 276:33196-33212(2001).
DB EMBL: M98457; AF77072.1;
DB EMBL: AF209479; AAG13814.1;
DB EMBL: AF290886; AAK20997.1;
DB EMBL: AY042238; AAK91863.1;
DB HSSP: P32897; IECG;
DB InterPro: IPR001304; Lectin_C1;
DB Pfam: PF00059; Lectin_C1;

```


DB 117 KSKQVEITFEETLTKAAVGLPEKSKIOEITVELTKAAVGLPEKSKIOEITVQDLSPL 176
 161 KAAVE-----LPDQSKQOUIYVELTDEKTAPEELDRHCPKD 197
 DB 177 KAAVGDLPKSKIOEITVQLTKAAVGLPEKSKIOEITVELTKAAVEELDRHCPME 256
 198 WTEFGNCTYFMSNSORNMHNSVTAQCEVKAJLVV:KIAEPAUNYGLV:ISNKR:SNMGLS 257
 DB 237 WTEFGNCTYFMSNSORNMHNSITACQEVGAQLVYIKSAEENFQI:QSSPSNPTMGLS 256
 258 DLNCGTQWQWDGSPSLSPFORYNNSGPNNSGNEDECAEFGSGMWNKCDVONVYICK 317
 DB 297 DLNCGTQWQWDGSPSLSPFORYNNSGPNNSGNEDECAEFGSGMWNKCDVONVYICK 356
 418 FVA-CFPRD 325
 DB 457 SAASCODE 365

RESULT 8

Q96002 PRELIMINARY: PRT: 381 AA.
 ID Q96002: PRELIMINARY: PRT: 381 AA.
 DT 01-JUN-2002 (TREMblrel: 21, Created)
 DT 01-JUN-2002 (TREMblrel: 21, Last sequence update)
 DT 01-JUN-2002 (TREMblrel: 21, Last annotation update)
 DB Type 11 membrane protein CD209.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Cladista; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21819420; PubMed=11818574,
 RA Wu L., Bashitova A.A., Martin T.D., Villamide J., Mehlof F.,
 KA Chetov A.O., Unutma D., Pope M., Carrington M., Kewakhamri V.N.;
 RT "Rhesus macaque dendritic cells efficiently transmit primate
 RT lentiviruses independently of DC-SIGN".
 RI Proc. Natl. Acad. Sci. U.S.A. 99:1568-1573(2002).
 DR EMBL: AY040319; AAK74185.1;
 SO SEQUENCE 381 AA; 42955 MW; 973D7B0H42C91D49 CRC64.

Query Match 81.1%; Score 1375; DB 6; Length 381.
 Best Local Similarity 74.0%; Pred. No. 38-82;
 Matches 273; Conservative 23; Mismatches 25; Indels 48; Gaps 4;

QY 1 MSDSKPRVVOGLCL-----GCLHGALVQLLSFMLAGVLVAI 40
 DB 1 MSDSKPRVVOGLCL-----GCLHGALVQLLSFMLAGVLVAI 40
 QY 41 LVQVSKVSSLSQEGSEDDATAYONTOLKAAVGLSEKSKIOEITVQLTKAAVGLPE 100
 DB 57 LVQVSKVSSLSQEGSEDDATAYONTOLKAAVGLSEKSKIOEITVQLTKAAVGLPE 116
 QY 101 KSKIOEITVQLTKAAVGLPEKSKIOEITVELTKAAVGLPEKSKIOEITVEL 160
 DB 117 KSKIOEITVQLTKAAVGLPEKSKIOEITVELTKAAVGLPEKSKIOEITVEL 176
 QY 161 KAAVE-----LPDQSKQOUIYVELTDEKTAPEELDRHCPKD 197
 DB 177 KAAVGDLPKSKIOEITVQLTKAAVGLPEKSKIOEITVELTKAAVEELDRHCPME 256
 QY 198 WTEFGNCTYFMSNSORNMHNSVTAQCEVKAJLVV:KIAEPAUNYGLV:ISNKR:SNMGLS 257
 DB 237 WTEFGNCTYFMSNSORNMHNSITACQEVGAQLVYIKSAEENFQI:QSSPSNPTMGLS 256
 258 DLNCGTQWQWDGSPSLSPFORYNNSGPNNSGNEDECAEFGSGMWNKCDVONVYICK 317
 DB 297 DLNCGTQWQWDGSPSLSPFORYNNSGPNNSGNEDECAEFGSGMWNKCDVONVYICK 356
 418 FVA-CFPRD 325
 DB 457 SAASCODE 365

DB 357 SAASCODE 365

RESULT 9

Q96001 PRELIMINARY: PRT: 404 AA.
 ID Q96001: PRELIMINARY: PRT: 404 AA.
 DT 01-DEC-2001 (TREMblrel: 19, Created)
 DT 01-DEC-2001 (TREMblrel: 19, Last sequence update)
 DT 01-MAR-2002 (TREMblrel: 20, Last annotation update)
 DE MDC-SIGNIB type I isoform.
 GN CD209.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Cladista; Vertebrata; Eu
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; H
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=1137487;
 RA Mummidi S., Catano G., Lam L., Hoelle A., Jelles V., J
 FA Jimenez F., Ahuja S.S., Ahuja S.K.;
 RT "Extensive Repertoire of Membrane-bound and Soluble Ig
 RT specific ICAM-3-grabbing Neutralizin 1 (DC-SIGN) and
 RT Isoforms: Inter-Individual Variation in Expression of
 RT Transcripts".
 RL J. Biol. Chem. 276:33196-33212(2001).
 DR EMBL: AY042224; AAK91854.1;
 DR InterPro: IPR001304; LectIn_C.
 DR Pfam: PF00456; LectIn_C_1.
 DR PROSITE: PS00615; G-TYPE-LECTIN_1; UNKNOWN_1.
 DR PROSITE: PS00411; C-TYPE-LECTIN_2; UNKNOWN_1.
 SO SEQUENCE 404 AA; 45570 MW; DD0570727D90C30 CRC64.

Query Match 80.8%; Score 1371; DB 4; Length
 Best Local Similarity 72.2%; Pred. No. 5-98-82;
 Matches 273; Conservative 23; Mismatches 23; Indels 48; Gaps 4;

QY 4 SKEPVQGLG-----LLOGLHGALVQLLSFMLAGVLVAI 40
 DB 15 SEEDLRGIGFPRQTRGYKSLAGTDLHPLVQLLSFTLLAG---LTVY 40
 QY 55 USKQALVYNTLTKAAVGLSEKSKIOEITVQLTKAAVGLPEKSKIOEITVEL 114
 DB 71 USKQALVYNTLTKAAVGLSEKSKIOEITVQLTKAAVGLPEKSKIOEITVEL 130
 QY 115 KAAVGLPEKSKIOEITVELTKAAVGLPEKSKIOEITVELTKAAVGLPEKSKIOEITVEL 160
 DB 131 KAAVGLPEKSKIOEITVELTKAAVGLPEKSKIOEITVELTKAAVGLPEKSKIOEITVEL 176
 QY 129 ELYOELTRKAAVGLPEKSKIOEITVELTKAAVGLPEKSKIOEITVELTKAAVGLPEKSKIOEITVEL 192
 DB 191 ELYOELTRKAAVGLPEKSKIOEITVELTKAAVGLPEKSKIOEITVELTKAAVGLPEKSKIOEITVEL 208
 QY 189 RLCHRCPKDMPFGQNCYFMSNSORNMHNSVTAQCEVKAJLVV:KIAEPAUNYGLV:ISNKR:SNMGLS 257
 DB 251 RLCHRCPKDMPFGQNCYFMSNSORNMHNSVTAQCEVKAJLVV:KIAEPAUNYGLV:ISNKR:SNMGLS 256
 QY 249 NPFSGMGLSTNDRFTWVWVPSN:SPSGRYNNSGPNNSGNEDECAEFGSGMWNKCDVONVYICK 317
 DB 311 NPFSGMGLSTNDRFTWVWVPSN:SPSGRYNNSGPNNSGNEDECAEFGSGMWNKCDVONVYICK 356
 QY 309 VQVWYICKKFAA-CFPRD 325
 DB 371 LAKFWICKKSAASGRDE 368
 RESULT 10
 Q96008 PRELIMINARY: PRT: 398 AA.
 ID Q96008: PRELIMINARY: PRT: 398 AA.
 DT 01-DEC-2001 (TREMblrel: 19, Created)
 DT 01-DEC-2001 (TREMblrel: 19, Last sequence update)
 DT 01-MAR-2002 (TREMblrel: 20, Last annotation update)

MD-STGNIA type II isoform.
 G6209.
 OS Homo sapiens (Human).
 NC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 NC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NC NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=11337487;
 RA Mummich S., Catano G., Lam L., Hoefle A., Telles V., Hoegun K.,
 RA Jimenez F., Ahuja S.S., Ahuja S.K.;
 RT "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
 RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-STG1) and DC-STG2
 RT Isoforms: Inter-individual Variation in Expression of DC-STG1
 RT Transcripts";
 RT J. Biol. Chem. 276:33196-33212(2001)
 DR EMBL: AY042222; AAK91847.1; C.
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF00059; Lectin_C_1.
 DR PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
 DR PROSITE: PS50041; C_TYPE_LECTIN_2.
 SO SEQUENCE 398 AA; 45031 MW; D925764886701D2E CRC64;

Query Match 80.78; Score 1369.5; DB 4; Length 498;
 Best Local Similarity 70.44; Pred. No 7; 3e-82;
 Matches 276; Conservative 20; Mismatches 19; Indels 77; Gaps 5;

QY 1 MSDSKEPVOQIGL-----GCLGHGALVLDLSPMLAGVIAI 40
 Db 1 MSDSKEPVOQIGLLEEGRLGLGFRGTGKSLAGLGGPLVLDLSTLLAG-----L 56
 QY 41 LVQSKVPSLSQPSQSDATYQNLTLQKAAGELSKSLQELIYQELTQKAAGELPE 100
 Db 57 LVQSKVPSLSQPSQSDATYQNLTLQKAAGELSKSLQELIYQELTQKAAGELPE 116
 QY 101 KSKLOEIYQELTRKAAGELPEKSL----- 127
 Db 117 KSKLOEIYQELTRKAAGELPEKSLQELIYQELTQKAAGELPEKSKQELIYQELTL 176
 QY 128 -----QELIYQELTRKAAGELPEKSLQELIYQELTQKAAGELPEKSKQELIYQELTL 174
 Db 177 KAAVGEIPEKSKQELIYQELTRKAAGELPEKSKQELIYQELTQKAAGELPEKSKQ 246
 QY 175 QIQGHTLTKTAFFRLCRHCPKMTFFQGNCTYMSNSQPMHNSVACQVAGLVVTK 244
 Db 237 EIQGHTLTKKAVERLCHPCWMTFFQGNCTYMSNSQPMHNSVACQVAGLVVTK 246
 QY 245 AEEVFLQLOTSRNSPESMMTSPINQGTWQWVDSPLSPSPQVWNSPEPNNSGNEP 294
 Db 297 AEE-----QSSRSNFTWMTSPINQGTWQWVDSPLSPSPQVWNSPEPNNSGNEP 350
 QY 295 AEFSSQMNWNTLVNWTCKKPAACQPHDE 325
 Db 351 AEFSSQMNWNTLVNWTCKKPAACQPHDE 382

RESULT 11
 Q95J96 PRELIMINARY; PRT; 381 AA.
 AC Q95J96;
 DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, last annotation update)
 DT Dendritic cell-specific ICAM-3 grabbing nonintegrin.
 OS Macaca mulatta (Pig tailed macaque)
 NC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 NC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 NC Cercopithecoidea; Macaca.
 NC NCB1_TaxID=9545;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21465051; PubMed 11581196;
 RA Baribaud F., Pahlmann S., Sparwasser T., Kimata M.T., Choi Y.K.,

Hagarty R.S., Ahmad N., Macfarlan T., Edwards T.G., Leslie G.J.,
 RA Atkinson J., Reihart T.A., Kimata M.T., Littman D.R., Hoxie J.A.,
 RA Davis R.W.;
 RT "Pig-tailed and rhesus dendritic cell-specific ICAM-3-grabbing nonintegrin,
 RT pig-tailed macaque, and murine dc-sign";
 RT J. Virol. 75:10281-10289(2001);
 DR EMBL: AF343727; AAL14428.1; C.
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF00059; Lectin_C_1.
 DR PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
 DR PROSITE: PS50041; C_TYPE_LECTIN_2; 1.
 SO SEQUENCE 491 AA; 42951 MW; C4F6E2E3D454B74A CRC64;

Query Match 80.58; Score 1465; DB 6; Length 481;
 Best Local Similarity 71.48; Pred. No 1; 4e-81;
 Matches 271; Conservative 23; Mismatches 27; Indels 48; Gaps 4;

QY 1 MSDSKEPVOQIGL-----GCLGHGALVLDLSPMLAGVIAI 40
 Db 1 MSDSKEPVOQIGLLEEGRLGLGFRGTGKSLAGLGGPLVLDLSTLLAG-----L 56
 QY 41 LVQSKVPSLSQPSQSDATYQNLTLQKAAGELSKSLQELIYQELTQKAAGELPE 100
 Db 57 LVQSKVPSLSQPSQSDATYQNLTLQKAAGELSKSLQELIYQELTQKAAGELPE 116
 QY 101 KSKLOEIYQELTRKAAGELPEKSLQELIYQELTQKAAGELPEKSKQELIYQELTL 160
 Db 117 KSKLOEIYQELTRKAAGELPEKSLQELIYQELTQKAAGELPEKSKQELIYQELTL 176
 QY 161 KAAVSE-----LPQSKQOQIYQELTDLKTAFFRLCRHCPK 197
 Db 177 KAAVGEIPEKSKQELIYQELTRKAAGELPEKSKQELIYQELTQKAAGELPEKSKQ 246
 QY 198 WPEFGNCTYMSNSQPMHNSVACQVAGLVVTKTAFFRLCRHCPKMTFFQGNCTYMS 257
 Db 237 WPEFGNCTYMSNSQPMHNSVACQVAGLVVTKTAFFRLCRHCPKMTFFQGNCTYMS 296
 QY 298 DINDGTWQWVDSPLSPSPQVWNSPEPNNSGNEPACAEPSQGMWQNDQVWNTCK 317
 Db 247 DINDGTWQWVDSPLSPSPQVWNSPEPNNSGNEPACAEPSQGMWQNDQVWNTCK 356
 QY 318 PAA-CPHDE 325
 Db 357 SAASCSQDE 365

RESULT 12
 Q95J96 PRELIMINARY; PRT; 404 AA.
 AC Q95J96;
 DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, last annotation update)
 DT Dendritic cell-specific ICAM-3 grabbing non-integrin.
 OS Macaca mulatta (Rhesus macaque)
 NC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 NC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 NC Cercopithecoidea; Macaca.
 NC NCB1_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=11581196;
 RA Baribaud F., Pahlmann S., Sparwasser T., Kimata M.T., Choi Y.K.,
 RA Davis R.W.;
 RT "Pig-tailed and rhesus dendritic cell-specific ICAM-3-grabbing nonintegrin,
 RT pig-tailed macaque, and murine dc-sign";
 RT J. Virol. 75:10281-10289(2001);
 DR EMBL: AF343727; AAL14428.1; C.
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF00059; Lectin_C_1.
 DR PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.


```

QY 135 TPLKAAGVLPKSKTQETIYQELTPLKAAGVLPKPSKQQTQYQELTTPKTAFFELCPHG 194
DB 173 TRKAAGVGLPEKSKQETIYQELTPLKAAGVGLPEKSKQETIYQELTOLKAAGVGLCPHG 232
QY 195 PKTWTFQGNWYMSNSQPMWHSVLAQGEVPAQLVLIKTAEDNPLQLOTSNPSRPSWM 254
DB 233 PWMTTFQGNWYMSNSQPMWHSVLAQGEVPAQLVLIKTAEDNPLQLOTSNPSRPSWM 292
QY 255 GLSDNDEGTWQWVDCSPITSPSFRWNSGCEFNNSGNEDEAFSGSGMNDKCDVNTW 314
DB 293 GLSDNDEGTWQWVDCSPITSPSFRWNSGCEFNNSGNEDEAFSGSGMNDKCDVNTW 352
QY 415 CKKPAA-CFRDE 325
DB 453 CKKSAASCSRDE 364

```

```

RESULT 15
Q960P9 PRELIMINARY: PRT: 334 AA.
AC Q960P9;
DI 01-DEC-2001 (TTEMblrel: 19, Created)
DI 01-DEC-2001 (TTEMblrel: 19, Last sequence update)
DI 01-MAR-2002 (TTEMblrel: 20, Last annotation update)
DE SMC-SIGNB type II isoform.
GN C0209.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN 111
RP SEQUENCE FROM N.A.
RX PubMed-11337487;
KA Mumenthal S., Catalan G., Lam I., Hoofnagle A., Telles V., Beatty K.,
KA Jimenez F., Ahuja S.S., Ahuja S.K.;
RT "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
RT Isoforms, Inter-individual Variation in Expression of DC-SIGN
RT Transcripts."
PL J. Biol. Chem. 276:33196-33212(2001).
DR EMBL: AY042231; AAK91856.1; -.
DR IntolPro: IPR001304; Lectin_G.
DR Pfam: PF000059; Lectin_G.
DR PROSITE: PS00615; G_TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE: PS50041; G_TYPE_LECTIN_2; -.
SV SOURCE 334 AA: 37843 MW: 86796pC4111C86BP CMC64;

```

```

Query Match 76.8%; Score 1003; DB 4; Length 334;
Rest Local Similarity 78.0%; Pred. No. 1,3e-77;
Matches 252; Conservative 25; Mismatches 26; Indels 20; Gaps 2;

```

```

QY 4 SKKRWQULDLGALVLDLSPMLAGVLAILEVQSKVPSLSQEDSPQDATTQ 63
DB 15 SPEQRLGLQFQHTRGYKSLA-----VSKVPSISQEDSPQDATTQ 55
QY 64 NLTQLKAAGVGLPEKSKQETIYQELTOLKAAGVGLPEKSKQETIYQELTOLKAAGVGLPE 123
DB 56 NLTQLKAAGVGLPEKSKQETIYQELTOLKAAGVGLPEKSKQETIYQELTOLKAAGVGLPE 115
QY 124 KSKLOETIYQELTOLKAAGVGLPEKSKQETIYQELTOLKAAGVGLPEKSKQETIYQELTOL 183
DB 116 KSKLOETIYQELTOLKAAGVGLPEKSKQETIYQELTOLKAAGVGLPEKSKQETIYQELTOL 175
QY 184 KLAFFELCPHGCPKDWTFPQGNWYMSNSQPMWHSVLAQGEVPAQLVLIKTAEDNPLQLOTS 243
DB 176 KLAFFELCPHGCPKDWTFPQGNWYMSNSQPMWHSVLAQGEVPAQLVLIKTAEDNPLQLOTS 235
QY 244 QTSRSNRFSSMGLSDNDEGTWQWVDCSPITSPSFRWNSGCEFNNSGNEDEAFSGSGMND 303
DB 236 QTSRSNRFSSMGLSDNDEGTWQWVDCSPITSPSFRWNSGCEFNNSGNEDEAFSGSGMND 295
QY 404 DNRCDVNDVWICKKPAA-CFRDE 325
DB 453 DNRCDVNDVWICKKPAA-CFRDE 364

```

DB 296 DNRCDVNDVWICKKPAA-CFRDE 318

Search completed: December 7, 2002, 11:05:43
Job Time: 91.600s


```

1 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
2 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
3 US-09-864-761-7350

```

Query Match 9.78; Score 153.2; DB 10; Length 592;
Best local similarity 98.1%; Prod No 6,9e+6;
Matches 15%, Conservative 0; Mismatches 3; Indels 0; Gaps 0;

DY 825 CATTGAACTCCCTGGCGGCAATGTGCCTAAGCACTGGACATCTCTCAGGAAGACTGTT 884
||| |||||||
DB 144 CAAAGAAAGTCTTCCTGGCGGCAATGTGCCTAAGCACTGGACATCTCTCAGGAAGACTGTT 203
||| |||||||

DY 885 ACTTGATCTCTTAATTCCTCAAGAGCAAGATCTGGACAGCATCTGGACAGCAAGACTCA 944
||| |||||||
DB 204 AATTCATCTCTTAATTCCTCAAGAGCAAGATCTGGACAGCATCTGGACAGCAAGACTCA 263
||| |||||||

DY 945 GGAGCGCACTCTGCTCTCAATTCATAAAAGCTGGCTCAAGCAAGAG 982
||| |||||||
DB 264 GGAGCGCACTCTGCTCTCAATTCATAAAAGCTGGCTCAAGCAAGAG 301
||| |||||||

RESULT 3
US-09-864-761-24972
Sequence 24972 Application US-09864761
Patent No. US2002048763A1
GENERAL INFORMATION:
APPLICANT: Feng, Sharron G.
APPLICANT: Park, David P.
APPLICANT: Ilanetz, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENEhme EXPRESSED SINGLE EXON NUCLEOTIC ACID PROBES USEFUL FO
FILE REFERENCE: Neomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
PENDING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,356
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: CR 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 00/246,454
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PG/9501/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PG/9501/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PG/9501/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PG/9501/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PG/9501/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PG/9501/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PG/9501/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PG/9501/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PG/9501/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PG/9501/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

[illegible]

PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/084508
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089653
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089908

Query Match
 Best Local Similarity 52.78; Score 64.2; EB 12; Length 577;
 Matches 188; Conservative 0; Mismatches 163; Indels 7; Gaps 2;

QY 777 AGTCAACACAGACAGAAATCTACAGAGAGCTGACCGATTGAAACATCTGATTCAGCCG 836
 DB 272 AGTTACACTCTGACAGATTTTCACAGAGCTCTCTGCTACATTAATGAGATTCAG 331
 QY 837 TGTGCGCCAGCTGCTCCAGAGCTGAGATTTCTTTCAGAACTCTTACTATCTGA 896
 DB 332 TCAAGAAATGTTGTCATTCAGTTCAGGATTAATTTCAATCAGCTGCTCTTTCTA 391
 QY 847 ACTTCAGTCAATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 956
 DB 342 CTGACACTATTTCTGCGCGCTTAAAGTTAAAGAACTCTGACCGATGCGGCTACCT 451
 QY 957 TCGTAATCAAACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1016
 DB 452 TGTATATCACTGACAGAGAGAGAGAGATTCCTTCTTCAAGAACTTAAATGAGAG 511
 QY 1017 GCTTCTCTGATGAGAGATTTTCAAGATTTAAATCAAGAGAGAGAGAGAGAGAG 1076
 DB 512 AGTTTCTTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 568
 QY 1077 GCTTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1133
 DB 569 GCACAGCTTTTACA---AAGTCTCTGAGCTTTCTGATGATGATGATGATGATGAT 622

RESULT 9
 US-09-831-481-995/C
 Sequence 995; Application US/09833381
 Patent No. US20020132090A1
 GENERAL INFORMATION:

APPLICANT: Robinson, Keith E
 TITLE OF INVENTION: NO. 0220020132090A1 Nucleic Acid and Protein Homologs
 FILE REFERENCE: 5800-119
 CURRENT APPLICATION NUMBER: US/09/833,381
 PRIOR FILING DATE: 2001-04-11
 PRIOR APPLICATION NUMBER: 09/516,448
 NUMBER OF SEQ ID NOS: 2050
 SOFTWARE: FASTSEQ for Windows Version 3.0
 SEQ ID NO 995
 LENGTH: 528
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (1)..(528)
 OTHER INFORMATION: n = A,T,C or G
 US-09-831-481-995

Query Match
 Best Local Similarity 3.68; Score 56; DB 10; Length 528;
 Matches 147; Conservative 0; Mismatches 130; Indels 3; Gaps 1;

QY 807 TGACGATTTTAAACATCTGATGATGATGATGATGATGATGATGATGATGATGAT 866
 DB 355 TGACGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 926
 QY 867 TCTTCAAGAGAACTGTTACTTATGCTAATCTGATGATGATGATGATGATGATGAT 926
 DB 295 GATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926
 QY 927 GCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 966

DB 235 AGACTGTTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926
 QY 987 TCTTCAAGAGAACTGTTACTTATGCTAATCTGATGATGATGATGATGATGATGAT 926
 DB 175 TCATGAG 926
 QY 1047 ATCTGAG 1086
 DB 118 AAGCTAAAG 79

RESULT 10
 US-09-880-107-3731
 Sequence 3731; Application US/09880107
 Patent No. US20020142981A1
 GENERAL INFORMATION:

APPLICANT: Horne, Daniel T.
 APPLICANT: Scherf, Uwe
 APPLICANT: Gene Logic, Inc.
 TITLE OF INVENTION: Gene Expression Profiles in Liver Can-
 FILE REFERENCE: 44921-5028-WO
 CURRENT APPLICATION NUMBER: US/09/880,107
 PRIOR FILING DATE: 2001-06-14
 PRIOR APPLICATION NUMBER: US 60/211,379
 PRIOR FILING DATE: 2000-06-14
 PRIOR APPLICATION NUMBER: US 60/217,054
 PRIOR FILING DATE: 2000-10-02
 NUMBER OF SEQ ID NOS: 3950
 SOFTWARE: Patent In Ver. 2.1
 SEQ ID NO 3731
 LENGTH: 1300
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: Genbank Accession No. US20020142981A1
 US-09-880-107-3731

Query Match
 Best Local Similarity 3.44; Score 52; DB 10; Length 1
 Matches 91; Conservative 0; Mismatches 65; Indels 0;

QY 847 CTGCTCCAG 926
 DB 646 CTGCTCCAG 926
 QY 907 GACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926
 DB 706 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926
 QY 967 AAGCTGAG 1002
 DB 766 CTGCTGAG 801

RESULT 11
 US-09-880-107-2235
 Sequence 2235; Application US/09880107
 Patent No. US20020142981A1
 GENERAL INFORMATION:

APPLICANT: Horne, Daniel T.
 APPLICANT: Scherf, Uwe
 APPLICANT: Gene Logic, Inc.
 TITLE OF INVENTION: Gene Expression Profiles in Liver Can-
 FILE REFERENCE: 44921-5028-WO
 CURRENT APPLICATION NUMBER: US/09/880,107
 PRIOR FILING DATE: 2001-06-14
 PRIOR APPLICATION NUMBER: US 60/211,379
 PRIOR FILING DATE: 2000-06-14
 PRIOR APPLICATION NUMBER: US 60/217,054
 PRIOR FILING DATE: 2000-10-02

[illegible]

1
2
3


```

RESULT 4
US-09-517-605-1
: Sequence 11, Application US/09517605
: Patent No. 6391567
: GENERAL INFORMATION:
: APPLICANT: Littman, Dan R.
: APPLICANT: Kwon, Douglas S.
: APPLICANT: Geljenbeck, Tracy
: APPLICANT: van Kooyk, Yvette
: TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
: FILE REFERENCE: 1049-1-017
: CURRENT APPLICATION NUMBER: US/09/517,605
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: Patent In Ver. 2.0
: SEQ ID NO 1
: LENGTH: 1312
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (42)..(1253)
US-09-517-605-1

```

```

Query Match 43.2% Score 666.4; DB 4; Length 1312;
Best Local Similarity 88.8%; Prid. No. 3,7e-169;
Matches 720; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

```

```

QY 424 TCAGAACAAATCCAGAGAGAGATCTACCGAACTGAGCACTTAAGCTAGCT 483
DB 383 TCCAGAGAAATCTAGCTAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 442
QY 484 GGGTACAGCTTCCAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAGAG 543
DB 443 GGGTACAGCTTCCAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAGAG 502
QY 544 GGGTACAGCTTCCAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAGAG 603
DB 503 GGGTACAGCTTCCAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAGAG 562
QY 604 CGGCTGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAGAG 663
DB 563 CGGCTGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAGAG 622
QY 664 CGGCTGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAGAG 723
DB 623 CGGCTGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAGAG 682
QY 724 GATCTACAGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 783
DB 683 GATCTACAGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 742
QY 784 GATCTACAGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 843
DB 743 GATCTACAGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 802
QY 844 GATCTACAGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 903
DB 803 GATCTACAGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 862
QY 904 GATCTACAGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 963
DB 863 GATCTACAGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 922
QY 964 GATCTACAGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 1023
DB 923 GATCTACAGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 982
QY 1024 GATCTACAGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 1083
DB 983 GATCTACAGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 1042

```

```

QY 1084 TCTGTCAACCAATCTAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 1143
DB 1043 TCTGTCAACCAATCTAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 1102
QY 1144 AGAGTGTCCGAGATTTAGCTAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 1203
DB 1103 AGAGTGTCCGAGATTTAGCTAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 1162
QY 1204 GATCTACAGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 1263
DB 1163 GATCTACAGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 1222

```

```

RESULT 4
US-09-591-435-11
: Sequence 11, Application US/09591435
: Patent No. 6280953
: GENERAL INFORMATION:
: APPLICANT: MESSIER, WALTER
: APPLICANT: STRELA, JAMES M
: TITLE OF INVENTION: METHODS TO IDENTIFY POLYNUCLEOTIDE AN
: TITLE OF INVENTION: SEQUENCES WHICH MAY BE ASSOCIATED W
: FILE REFERENCE: GPN0.200.2
: CURRENT APPLICATION NUMBER: US/09/591,435
: PRIOR FILING DATE: 2000-06-09
: PRIOR APPLICATION NUMBER: 09/591,435
: PRIOR FILING DATE: 2000-06-09
: PRIOR APPLICATION NUMBER: 09/240,945
: PRIOR FILING DATE: 1999-01-29
: PRIOR APPLICATION NUMBER: 60/607,264
: PRIOR FILING DATE: 1998-01-30
: PRIOR APPLICATION NUMBER: 60/094,987
: PRIOR FILING DATE: 1998-09-02
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: Patent In Ver. 2.0
: SEQ ID NO 11
: LENGTH: 1212
: TYPE: DNA
: ORGANISM: Gallus gallus
US-09-591-435-11

```

```

Query Match 42.7% Score 659; DB 4; Length
Best Local Similarity 88.8%; Prid. No. 3.9e-167;
Matches 716; Conservative 0; Mismatches 95; Ind

```

```

QY 424 TCAGAACAAATCCAGAGAGAGATCTACCGAACTGAGCACTTAAGCTAGCT 483
DB 383 TCCAGAGAAATCTAGCTAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 442
QY 484 GGGTACAGCTTCCAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAGAG 543
DB 443 GGGTACAGCTTCCAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAGAG 502
QY 544 GGGTACAGCTTCCAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAGAG 603
DB 503 GGGTACAGCTTCCAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAGAG 562
QY 604 CGGCTGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAGAG 663
DB 603 CGGCTGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAGAG 622
QY 664 CGGCTGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAGAG 723
DB 623 CGGCTGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAGAG 682
QY 724 GATCTACAGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 783
DB 723 GATCTACAGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 742
QY 784 GATCTACAGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 843
DB 783 GATCTACAGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 802
QY 844 GATCTACAGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 903
DB 843 GATCTACAGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 862
QY 904 GATCTACAGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 963
DB 903 GATCTACAGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 922
QY 964 GATCTACAGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 1023
DB 963 GATCTACAGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 982
QY 1024 GATCTACAGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 1083
DB 1023 GATCTACAGAGAGCTGAGAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 1042

```

	Query Match	42-48	Score 653.8	PR 4	Length 1212
	Rest Local Similarity	87.8%	Pred. No. 4,66-166		
	Matches 712	(conservative)	1	Mismatches 98	Indels 0
				Gaps	0
QY	424	TCAGAGCAATTCGAGCAACATGCTAAATTCATGACAGCTGACCTTAAAGCTGTACG	483		
146	142	TCAGAGCAAAATTAACGACATACATCATCTACATGAGAGCTGAGAGGCTGTACG	401		
QY	484	GAGTGAGCTTTATAGAGAAATTAATTTATAGAGATGATAGAGAGAGAGG	543		
146	402	GAGTGAGCTTTCGAGAGCAAACTTAACTATCATATATTAATAGAGAGCTGAGAG	441		
QY	544	GAGTGAGCTGAGCTGTCTATAGAGAAATTAATTTATAGAGAGATTAATAGAGATGAG	603		

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

```

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,145
REFERENCE/DOCKET NUMBER: ITTX-0493
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-0000
FAX: 512/418-0000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 501 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

```

Query Match	3.68;	Score 56;	DB 3;	Length 501;
Host Local Similarity	52.5%;	Prod. No. 8.8e-06;		
Matches 147;	Conservative	0;	Mismatches 130;	Indels 5;
			Gaps 1;	

QY	807	TACGGCATTTGGAAAGTATGTCATTTTGAACAGCGCTGTGGAGAGACAGTGTGCCAAGAGATGGAAAT	866
DB	68	TGAGTGAAGAGCAATATGGTATTTAGAAAAATATGGCATATGGTATGGTAAATTAATGTAAGT	127
QY	867	TCTTTCAGAGTAATCTGTATCTCATCTGTCTTAAGTCTCTATGTAAGTGAAGTGGACCTGGATCA	926
DB	128	CATTGGCTGCGACATCTCTACCGCATTTCTTACGACAGAGAAACCTTTGGAGGACGATGAGGC	187
QY	927	CTGCTGTCTTAAATGATGAAATGGAGGCTGAGTCTGGTCAATCTAAAGATGCTGAGAGATCAAGAT	986
DB	188	AGAACTGCTGTTCCAGATGGAGGGCTGCACTGCTGATGATTAATCTTAAAGCTGAGACAGAT	247
QY	987	TCTTAGAGCTGAGACCTTCAGACATCAACAGCTTCTCTGATGAGGACCTTCAAGCTTAA	1046
DB	248	TCATCAGCCGACGAGCTGAAATGACTCACTCTTCTTACTTCTCTGGTCTCTTTCGGAACGAC	304
QY	1047	ATCAGCAAGGACAGTGGCAATATGGTGGAGAGCGCTCAACTT	1086
DB	305	AAGCTAATGCTAATGCTCAATGATATGATGATAATCTGTTT	344

RESULT 7
 US-08-772-440-3
 Sequence 3, Application US/08772440
 Patent No. 6046158
 GENERAL INFORMATION:
 APPLICANT: Atizumi, Kiyoshi
 APPLICANT: Takashima, Akira
 TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
 TITLE OF INVENTION: LECTINS, DCTN-1 AND DCTN-2; COMPOSITIONS AND USES
 TITLE OF INVENTION: THEREOF
 NUMBER OF SEQUENCES: 42
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Arnold, White & Burkee
 STREET: P.O. Box 4433
 CITY: Houston
 STATE: Texas
 COUNTRY: USA
 ZIP: 77210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/772,440
 FILING DATE: CONCURRENTLY HERewith
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Parker, David L.
 REGISTRATION NUMBER: 32,165
 REFERENCE/DOCKET NUMBER: 07X0-493
 TELECOMMUNICATION INFORMATION:

```

: TELEPHONE: 512/418-3000
:
: TELEFAX: 512/474-7577
:
: INFORMATION FOR SEQ ID NO:
:
: SEQUENCE CHARACTERISTICS:
:
: LENGTH: 1227 base pairs
:
: TYPE: nucleic acid
:
: STRANDEDNESS: single
:
: TOPOLOGY: linear

```

Query Match	4.68;	Score	56;	DB	4;	Length	1
Best Local Similarity	52.5%;	Pred. No.	1.3e-05;				
Matches	147;	Conservative	0;	Mismatches	130;	Indel	

[illegible]

RESULT 8
 US-08-809-494A-1
 Sequence 1, Application US/088094A
 Patent No. 5962260
 GENERAL INFORMATION:
 APPLICANT: Sawamura, Tatsuya
 APPLICANT: Masaki, Tomoo
 TITLE OF INVENTION: Modified Low-Density Lipoprotein
 TITLE OF INVENTION: Receptor
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: McAlay Fisher Nissan Goldberg & Kiel
 STREET: 261 Madison Avenue
 CITY: New York
 STATE: NY
 COUNTRY: USA
 ZIP: 10016-2391
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.10
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/809-494A
 FILING DATE: 24-MAR-1997
 CLASSIFICATION: A35
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 6-321705
 FILING DATE: 30-NOV-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 7-214206
 FILING DATE: 31-JUL-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Goldberg, Jules E.
 REGISTRATION NUMBER: 24408
 REFERENCE/DOCKET NUMBER: JG-V-4303CT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212 986-4090

[illegible]

```

CORRESPONDENCE ADDRESS:
ADDRESSER: MCALRAY FISHOT Nissen Goldberg & Kiel
STREET: 261 Madison Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10016-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 88/073522_302
FILING DATE: 12-JUL-1993
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-421705
FILING DATE: 30-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-214206
FILING DATE: 31-JUL-1995
ATTORNEY/AGENT INFORMATION:
NAME: Goldberg, Jules E
REGISTRATION NUMBER: 24408
REFERENCE/TOKENET NUMBER: JC-VY-436JCT/D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 985-4990
TELEFAX: 212 918-3479
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1897 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bos taurus
TISSUE TYPE: Vascular endothelial cells
IMMEDIATE SOURCE:
LIBRARY: Bovine aortic endothelial cell cDNA
CLONE: pBOX-1
FEATURE:
NAME/KEY: polyA_site
LOCATION: 1895..1897
FEATURE:
NAME/KEY: misc_RNA
LOCATION: 1859..1864
OTHER FEATURE: "polyA signal"
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..34
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 848..1897
FEATURE:
NAME/KEY: CDS
LOCATION: 35..847
US-09-352-302-1
Query Match: 3.5%; Score 54.6; DB 4; Length 1897;
Best Local Similarity 48.8%; Fred No. 3.5e-05;
Matches 147; conservative 0; Mismatches 154; Indels 0; Gaps 0;
QY ACCGATACAGCGCATTTGAATACATGCATTGGAAACCCGTGTCGCCACTGTCCTCAAG 857
DB ACCGATACAGCGCATTTGAATACAGCGCATTTGAAGCGCGAACCAACTATTCAGTGCTTTCGCGAAG 461
UY ACTGGAATATCTTCGCAAGCAAACTGTACTCATGATGCTTAACCTCCAAAGCAACAGCGAG 917
DB ACTGGAATATCTTCGCAAGCAAACTGTACCAATTTGCTCTGGGCTCTTTTAATTTGGAAA 521

```


[illegible][illegible]

[illegible]

QY	104	107	QY	364	QY	424	QY	484	QY	544	QY	604	QY	664	QY	724	QY	786	QY	846	QY	906	QY	966	QY	1026	QY	1086	QY	1146	QY	1206	QY	1266	QY	1326	QY	1386	QY	1446	QY	1506	QY	1566	QY	1626	QY	1686	QY	1746	QY	1806	QY	1866	QY	1926	QY	1986	QY	2046	QY	2106	QY	2166	QY	2226	QY	2286	QY	2346	QY	2406	QY	2466	QY	2526	QY	2586	QY	2646	QY	2706	QY	2766	QY	2826	QY	2886	QY	2946	QY	3006	QY	3066	QY	3126	QY	3186	QY	3246	QY	3306	QY	3366	QY	3426	QY	3486	QY	3546	QY	3606	QY	3666	QY	3726	QY	3786	QY	3846	QY	3906	QY	3966	QY	4026	QY	4086	QY	4146	QY	4206	QY	4266	QY	4326	QY	4386	QY	4446	QY	4506	QY	4566	QY	4626	QY	4686	QY	4746	QY	4806	QY	4866	QY	4926	QY	4986	QY	5046	QY	5106	QY	5166	QY	5226	QY	5286	QY	5346	QY	5406	QY	5466	QY	5526	QY	5586	QY	5646	QY	5706	QY	5766	QY	5826	QY	5886	QY	5946	QY	6006	QY	6066	QY	6126	QY	6186	QY	6246	QY	6306	QY	6366	QY	6426	QY	6486	QY	6546	QY	6606	QY	6666	QY	6726	QY	6786	QY	6846	QY	6906	QY	6966	QY	7026	QY	7086	QY	7146	QY	7206	QY	7266	QY	7326	QY	7386	QY	7446	QY	7506	QY	7566	QY	7626	QY	7686	QY	7746	QY	7806	QY	7866	QY	7926	QY	7986	QY	8046	QY	8106	QY	8166	QY	8226	QY	8286	QY	8346	QY	8406	QY	8466	QY	8526	QY	8586	QY	8646	QY	8706	QY	8766	QY	8826	QY	8886	QY	8946	QY	9006	QY	9066	QY	9126	QY	9186	QY	9246	QY	9306	QY	9366	QY	9426	QY	9486	QY	9546	QY	9606	QY	9666	QY	9726	QY	9786	QY	9846	QY	9906	QY	9966	QY	10026	QY	10086	QY	10146	QY	10206	QY	10266	QY	10326	QY	10386	QY	10446	QY	10506	QY	10566	QY	10626	QY	10686	QY	10746	QY	10806	QY	10866	QY	10926	QY	10986	QY	11046	QY	11106	QY	11166	QY	11226	QY	11286	QY	11346	QY	11406	QY	11466	QY	11526	QY	11586	QY	11646	QY	11706	QY	11766	QY	11826	QY	11886	QY	11946	QY	12006	QY	12066	QY	12126	QY	12186	QY	12246	QY	12306	QY	12366	QY	12426	QY	12486	QY	12546	QY	12606	QY	12666	QY	12726	QY	12786	QY	12846	QY	12906	QY	12966	QY	13026	QY	13086	QY	13146	QY	13206	QY	13266	QY	13326	QY	13386	QY	13446	QY	13506	QY	13566	QY	13626	QY	13686	QY	13746	QY	13806	QY	13866	QY	13926	QY	13986	QY	14046	QY	14106	QY	14166	QY	14226	QY	14286	QY	14346	QY	14406	QY	14466	QY	14526	QY	14586	QY	14646	QY	14706	QY	14766	QY	14826	QY	14886	QY	14946	QY	15006	QY	15066	QY	15126	QY	15186	QY	15246
----	-----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------	----	-------

Accession	AY042326	GI:15281092
Version	AY042326.1	GI:15281092
Keywords		
Source		
Organism	Homo sapiens.	
Reference Authors	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata Mammalia; Eutheria; Primates; Catarrhini; Hominoidea i (bases 1 to 1083) Mummid, S., Catano, O., Lam, L., Hoeffle, A., Telles, Jlencez, F., Ahuja, S.S., and Ahuja, S.K. Extensive repertoire of membrane-bound and soluble cell specific TAM-activating molecules in 1 (the isoforms. Inter-individual variation in expressed transcripts	
Title		
Journal	J. Biol. Chem. 276 (45), 4169-4212 (2001)	
Medline	21413847	
PubMed	11337487	
Reference	2 (bases 1 to 1083)	
Authors	Mummid, S., and Ahuja, S.K.	
Title	Direct Submission	
Journal	Submitted (18-JUN-2001) Department of Medicine, Veterans Health Care System and University of Texas Center, 7703, Floyd Curl Drive, San Antonio, TX	
Features	Location:Qualifiers	
Source	1..1083	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/chromosome="19"	
	/map="19p13.4"	
	1..1083	
gene	/gene="CD209"	
	1..1083	
	/gene="CD209"	
	/note="Similar to the nucleotide sequen Genbank Accession Number AC008812; alt= /codon_start=-1 /product="SPC-SIGNIA TYPE II Isoform" /protein_id="AAK91851.1" /db_xref="GI:15281083"	
	/translation="MEQSKETROQJLDLYSKVSTED GEISPKRIQIYQELTQTKAAVETPEKSKQIYQEI YQELTQTKAAVETPEKSKQIYQELTQTKAAVETPEKSKQIYQEI ELPEKSKQIYQELTQTKAAVETPEKSKQIYQELTQTKAAVETPEKSKQIYQEI FOGNCYFENSQSRMHQSTTACKKVEQALVITSAFLN LNQGTQWQWQWQSPLEPSFQYQWNRGPPNVVCFEPAFI KSAQSKRDHQLHSPAPLNPVPPA"	
BASE COUNT	290 a 284 c 410 g 199 t	
ORIGIN		
Query Match	43.2% Score 665.4; DB 9; Ident	
Best Local Similarity	88.8% Ident. No. 1; Ident	
Matches 720, Conservative: 6, Mismatches 91, Indels		
424 TCAGAGCAATCCGACGACATATCTACTAGACCTGACCCAGCTTAA		
210 TCACAGAAATCTAAGCTGACAGATATATTAAGATGACCTGCTG		
444 GGTGAGCTCTCAGAGAAATCTAAGCTGACAGATATATTAAGATGACCTGCTG		
270 GGTGAGCTCTCAGAGAAATCTAAGCTGACAGATATATTAAGATGACCTGCTG		
544 GGCTGCAGTGGGTGAGTCTGACAGAGAAATCTAAGCTGACAGATATAT		
330 GGCTGCAGTGGGTGAGTCTGACAGAGAAATCTAAGCTGACAGATATAT		
604 GGCTGCAGTGGGTGAGTCTGACAGAGAAATCTAAGCTGACAGATATAT		
390 TCGGTGAAGGTGACAGTGGGTGAGTCTGACAGAGAAATCTAAGCTGACAG		
664 GGAATGAGTGGGTGAGTCTGACAGAGAAATCTAAGCTGACAGAGAAATCTAAG		
450 GGAATGAGTGGGTGAGTCTGACAGAGAAATCTAAGCTGACAGAGAAATCTAAG		

Db	762	CGGCTGCTGCTGGGAAAGGAAATCTCTCCAGGAAGCTGTACTTCATCTGTCAATGCTCA	821
Gy	964	GAGGAAATGATAGGATATGGGATCAAGGCTGGCTAGGAAAGTAGGAGGAGGAGGAGTCTGGTGGTAA	963
Db	822	CGGGAATCTGGACGAGCTCCATCAACCCCTGGCAAAAGAACTGGGGGCCAGGCTCGCTGTAAT	881
Gy	964	CAAAATCTGCTTAAGTACAGTACAACTCTTCTTAACTCTGACACTTCCAGGATCAACCTGCTCTG	1023
Db	882	CAAAAGCTGCTGAGGATATAGATCTCTCTATAGCTGAGCTCTCCAGGAATTAACCTGCTTCAAC	941
Gy	1024	CTGATCTGGGACTTTCTACACTTAATATAGGAAAGGAGCTGGTGGTAAATGGTGGCTCCACG	1083
Db	942	CTGGATATGGGACTTCTCGGATCTTAATATAGGAAAGGAGGAGCTGGTGGTAAATGGGAGTAA	1091
Gy	1084	TCTGTCACTCAAGCTTCCAGAGGCTACTGGAAAGTAGGAAACCCAAACATATGGCGGAAATTA	1143
Db	1082	CTCTGTGCGCAGGCTTCAGAGCAGTATTTGGGAAATAGATATAGAGCTTCAATAACTGTGGGAGGA	1061
Gy	1144	AAGATGTGGTAAATTTAGTGGGAGTGGTGGTGGAGGAAATATGGATCTCACTGTACCAATTA	1203
Db	1062	AAGATCTGGGAAATTTAGTGGTAAATGGTGGGAGGAGGAGCAAAATGTATCTTCCAAATG	1121
Gy	1204	CTGGAATCTGCAAAAGCCCGGACCTGCTCTG	1234
Db	1122	CTGATCTGCAAAAGTCCGAGCTGCTCTG	1152

RESULT 14			
AX039403			
LOCUS	AX039403	1215 bp	FNA
DEFINITION	Sequence 1 from Patent WO0063251.		
ACCESSION	AX039403		
VERSION	AX039403.1	GI:11229473	

ORGANISM	REFERENCE	AUTHORS	TITLE
<i>Homo sapiens</i>			
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;			
Mammalia: Eutheria: Primates: Catarrhini, Homnidae: Homo.			
1 (Phase 1 to 1215)			
Fiddler, C.G., Geijzenbeek, T. B., van Kooyk, Y., and Torensmma, R.			
Composition and method for modulating dendritic cell-t cell intera			
ction			

ORIGINAL	PATIENT	WORLDWIDE ID	AGE	SEX	DATE
Katholieke Universiteit Nijmegen (NL)	W0064251	A	F	2004	
FEATURES	LOCATION/QUALIFIERS				
SOURCE	1..1215	/organism="Homo sapiens"			
		/db_xref="taxon:9606"			
		/note="Human: DC-SIGN"			
BASE COUNT	315	a	321	c	349
ORIGIN					230

```
Query Match:      43.28; Score 665.4; PR 6; Length 1215.
Best Local Similarity: 88.88; Fred No. 1.3e-162;
Matches 720; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
```

QY 424 TCAGGAACTCCGACAGACGATATTATTCAGCAACTGCACAGTGAAGGTGGAGT 483
|| || ||| || | || || || || || || || || || || || || || || || || || || ||
Db 342 TCCAGAGAAATCTTAAGCTCGAGATAGATTATTCAGSAGCTGACCGCGCTGGAAGGCTGAGAGT 401

cy	484	GGTGGATTTCTTGAGAAATCTAAATCTGATGATTTTCAGAGCTGACCTACCTGAA	543
db	402	GGTGGATTTCTTGAGAAATCTAAATCTGATGATTTTCAGAGCTGACCTACCTGAA	461

[illegible]

5.22 TGGGCTGAAGCTGCTGATGATCTTCCAGAGAAATCTAAGCAGCAGAGAAATCTACCA 581

QY	664	GAAGTATACGCGGCGCTCAAGAGCTGCATGGGGGCGCATCTGGCGAGCAAAATTCAGAGCTGTAAGCA	723
Dp	582	CGAGATTACCTGGGCTCAAGAGCTGCATGGGGGCGCATCTGGCGAGCAAAATTCAGAGCTGTAAGCA	641
QY	724	GATGTAACTAAGGAGCTCAAGAGCTGCATGGGGGCGCATCTGGCGAGCAAAATTCAGAGCTGTAAGCA	783
Dp	642	GATGTAACTAAGGAGCTCAAGAGCTGCATGGGGGCGCATCTGGCGAGCAAAATTCAGAGCTGTAAGCA	701
QY	784	GGAGTAACTAAGGAGCTCAAGAGCTGCATGGGGGCGCATCTGGCGAGCAAAATTCAGAGCTGTAAGCA	843
Dp	702	GGAGTAACTAAGGAGCTCAAGAGCTGCATGGGGGCGCATCTGGCGAGCAAAATTCAGAGCTGTAAGCA	761
QY	844	CGAGTAACTAAGGAGCTCAAGAGCTGCATGGGGGCGCATCTGGCGAGCAAAATTCAGAGCTGTAAGCA	903
Dp	762	CGAGTAACTAAGGAGCTCAAGAGCTGCATGGGGGCGCATCTGGCGAGCAAAATTCAGAGCTGTAAGCA	821
QY	904	GGAGTAACTAAGGAGCTCAAGAGCTGCATGGGGGCGCATCTGGCGAGCAAAATTCAGAGCTGTAAGCA	963
Dp	822	GGAGTAACTAAGGAGCTCAAGAGCTGCATGGGGGCGCATCTGGCGAGCAAAATTCAGAGCTGTAAGCA	881
QY	964	CGAGTAACTAAGGAGCTCAAGAGCTGCATGGGGGCGCATCTGGCGAGCAAAATTCAGAGCTGTAAGCA	1023
Dp	882	CGAGTAACTAAGGAGCTCAAGAGCTGCATGGGGGCGCATCTGGCGAGCAAAATTCAGAGCTGTAAGCA	941
QY	1024	CGAGTAACTAAGGAGCTCAAGAGCTGCATGGGGGCGCATCTGGCGAGCAAAATTCAGAGCTGTAAGCA	1083
Dp	942	CGAGTAACTAAGGAGCTCAAGAGCTGCATGGGGGCGCATCTGGCGAGCAAAATTCAGAGCTGTAAGCA	1001
QY	1084	CGAGTAACTAAGGAGCTCAAGAGCTGCATGGGGGCGCATCTGGCGAGCAAAATTCAGAGCTGTAAGCA	1143
Dp	1002	CGAGTAACTAAGGAGCTCAAGAGCTGCATGGGGGCGCATCTGGCGAGCAAAATTCAGAGCTGTAAGCA	1061
QY	1144	CGAGTAACTAAGGAGCTCAAGAGCTGCATGGGGGCGCATCTGGCGAGCAAAATTCAGAGCTGTAAGCA	1203
Dp	1062	CGAGTAACTAAGGAGCTCAAGAGCTGCATGGGGGCGCATCTGGCGAGCAAAATTCAGAGCTGTAAGCA	1121
QY	1204	CGAGTAACTAAGGAGCTCAAGAGCTGCATGGGGGCGCATCTGGCGAGCAAAATTCAGAGCTGTAAGCA	1263
Dp	1122	CGAGTAACTAAGGAGCTCAAGAGCTGCATGGGGGCGCATCTGGCGAGCAAAATTCAGAGCTGTAAGCA	1181

RESULT 15					
AY042221	AY042221	1,215 bp	mpna	linear	181 59-AUG-2001
LOCUS					
DEFINITION	Homo sapiens MUC 5B1A type 1 isoform (CD209) mRNA, complete cds,				
ACCESSION	AY042221				
VERSION	AY042221.1				

NETWORK SOURCE	ORGANISM	REFERENCE	TITLE
Homo sapiens.			
Homo sapiens.			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.			
(Class 1 to 1215)			
Homodonta; Catarrhini; Lani, L., MacFie, A., Jellies, V., Bequm, K., Jimenez, F., Anuja, S.S., and Anuja, S.K.			
Extensive repertoire of membrane-bound and soluble dendritic			

JOURNAL: J Biol Chem. 276 (45), 33106-33112 (2001)
 MEDLINE: 2111847
 PUBMED: 11317487
 REFERENCE: 2 (bases 1 to 1215)
 AUTHORS: Mummidi, S. and Abuji, S. K.
 TITLE: Direct Submission

JOURNAL Submitted (18-JUN-2001) Department of Medicine, South Texas Veterans Health Care System and University of Texas Health Science Center, 7703, Floyd Curl Drive, San Antonio, TX 78229, USA

FEATURES LOCATION/QUALIFIERS
1. 1215
/organism="Homo sapiens"
/db_xref="taxon:9606"

...

Run on: December 7, 2002, 11:02:30 - Search time: 21 seconds
(without alignments)

251.369 Million cell updates/sec

Title: US-09-831-458A-12
1606

Sequence: 1 MSDSKEPRVQQLGLLGLIGH PCDDVDNYWICKKPPACFRDE 325

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched. 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters:	10394
--	-------

Maximum DB seq length: 200000000000

Post-processing:	Minimum Match	0%
	Maximum Match	100%

Listing first 45 summaries

```
Database : published_Applications_AA:
1 /cvm3_6/rotdata/3/published
```

1	/cgn2_2/prodata/2/pubpa/1/ISO1_NEW_PUB pep *
2	/cgn2_2/prodata/2/pubpa/1/ISO6_NEW_PUB pep *
3	/cgn2_2/prodata/2/pubpa/1/ISO6_NEW_PUB pep *
4	/cgn2_2/prodata/2/pubpa/1/ISO6_FUBCOMB pep *
5	/cgn2_2/prodata/2/pubpa/1/ISO7_FUBCOMB pep *
6	/cgn2_2/prodata/2/pubpa/1/ISO7_FUBCOMB pep *
7	/cgn2_2/prodata/2/pubpa/1/ISO8_FUBCOMB pep *
8	/cgn2_2/prodata/2/pubpa/1/ISO8_FUBCOMB pep *
9	/cgn2_2/prodata/2/pubpa/1/ISO8_NEW_PUB pep *
10	/cgn2_2/prodata/2/pubpa/1/ISO9_FUBCOMB pep *
11	/cgn2_2/prodata/2/pubpa/1/ISO9_NEW_PUB pep *
12	/cgn2_2/prodata/2/pubpa/1/ISO10_FUBCOMB pep *
13	/cgn2_2/prodata/2/pubpa/1/ISO10_NEW_PUB pep *
14	/cgn2_2/prodata/2/pubpa/1/ISO11_FUBCOMB pep *

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	319.5	18.8	316	9	US-09-802-802-4	Sequence 4, Appl
2	305.2	18.0	273	9	US-09-802-802-10	Sequence 20, Appl
3	289.5	17.1	293	9	US-09-878-295A-231	Sequence 231, Appl
4	289.6	17.1	294	9	US-09-878-097-431	Sequence 231, Appl
5	289.5	17.1	294	9	US-09-878-142A-231	Sequence 231, Appl
6	284.6	16.7	50	10	US-09-804-761-40945	Sequence 40945
7	274.5	16.2	323	10	US-09-764-870-279	Sequence 279, Appl
8	273.5	16.1	219	9	US-09-805-291A-377	Sequence 377, Appl
9	273.5	16.1	219	9	US-09-804-420-377	Sequence 377, Appl
10	273.5	16.1	219	9	US-09-909-088B-377	Sequence 24, Appl
11	273.5	16.1	219	12	US-10-062-588-24	Sequence 24, Appl
12	271.5	16.0	291	9	US-09-862-802-5	Sequence 6, Appl
13	267.5	15.8	287	9	US-09-862-802-5	Sequence 6, Appl
14	263.5	15.5	182	12	US-10-040-466-2	Sequence 2, Appl
15	263.5	15.5	182	12	US-10-040-466-2	Sequence 2, Appl
16	246	14.5	273	10	US-09-796-856-47	Sequence 47, Appl
17	244.6	13.8	403	9	US-09-878-759-142	Sequence 2, Appl
18	232	13.7	237	9	US-09-862-802-49	Sequence 49, Appl
19	232	13.7	237	9	US-09-862-759-49	Sequence 49, Appl

20	232	243	10	US-09-764-870-463	10	US-09-764-870-463	10	US-09-764-870-463	10	US-09-764-870-463
21	232	243	10	US-09-764-870-463	10	US-09-764-870-463	10	US-09-764-870-463	10	US-09-764-870-463
22	230	240	10	US-09-745-763-218	10	US-09-745-763-218	10	US-09-745-763-218	10	US-09-745-763-218
23	230	240	10	US-09-745-763-218	10	US-09-745-763-218	10	US-09-745-763-218	10	US-09-745-763-218
24	226.5	236.5	9	US-09-862-802-6	9	US-09-862-802-6	9	US-09-862-802-6	9	US-09-862-802-6
25	210.5	220.5	9	US-09-870-759-95	9	US-09-870-759-95	9	US-09-870-759-95	9	US-09-870-759-95
26	206	216	10	US-09-644-807-439	10	US-09-644-807-439	10	US-09-644-807-439	10	US-09-644-807-439
27	202.5	212.5	10	US-09-764-870-463	10	US-09-764-870-463	10	US-09-764-870-463	10	US-09-764-870-463
28	202.5	212.5	10	US-09-764-870-463	10	US-09-764-870-463	10	US-09-764-870-463	10	US-09-764-870-463
29	189	199	10	US-09-764-870-463	10	US-09-764-870-463	10	US-09-764-870-463	10	US-09-764-870-463
30	189	199	10	US-09-764-870-463	10	US-09-764-870-463	10	US-09-764-870-463	10	US-09-764-870-463
31	188.5	198.5	11	US-10-052-598-684	11	US-10-052-598-684	11	US-10-052-598-684	11	US-10-052-598-684
32	186	196	12	US-09-764-870-463	12	US-09-764-870-463	12	US-09-764-870-463	12	US-09-764-870-463
33	186	196	10	US-09-811-679-110	10	US-09-811-679-110	10	US-09-811-679-110	10	US-09-811-679-110
34	183.5	193.5	9	US-09-912-598-357	9	US-09-912-598-357	9	US-09-912-598-357	9	US-09-912-598-357
35	183.5	193.5	9	US-09-989-2938-357	9	US-09-989-2938-357	9	US-09-989-2938-357	9	US-09-989-2938-357
36	183.5	193.5	10	US-09-989-722-357	10	US-09-989-722-357	10	US-09-989-722-357	10	US-09-989-722-357
37	183.5	193.5	10	US-09-989-722-357	10	US-09-989-722-357	10	US-09-989-722-357	10	US-09-989-722-357
38	183.5	193.5	10	US-09-989-722-357	10	US-09-989-722-357	10	US-09-989-722-357	10	US-09-989-722-357
39	183.5	193.5	10	US-09-989-722-357	10	US-09-989-722-357	10	US-09-989-722-357	10	US-09-989-722-357
40	183.5	193.5	10	US-09-989-722-357	10	US-09-989-722-357	10	US-09-989-722-357	10	US-09-989-722-357
41	183.5	193.5	10	US-09-989-722-357	10	US-09-989-722-357	10	US-09-989-722-357	10	US-09-989-722-357
42	183.5	193.5	10	US-09-989-722-357	10	US-09-989-722-357	10	US-09-989-722-357	10	US-09-989-722-357
43	183.5	193.5	10	US-09-989-722-357	10	US-09-989-722-357	10	US-09-989-722-357	10	US-09-989-722-357
44	183.5	193.5	10	US-09-989-722-357	10	US-09-989-722-357	10	US-09-989-722-357	10	US-09-989-722-357
45	183.5	193.5	10	US-09-989-722-357	10	US-09-989-722-357	10	US-09-989-722-357	10	US-09-989-722-357

ALLEGMENTS

```

RESULT 1
US-09-862-802-4
? Sequence 4, Application US/0464286/2
? Patent No. US20020165346A1
?
? GENERAL INFORMATION: -PLONG CORPORATION
? APPLICANT: SCHRÖDING, MAAMALLAN MEMBRANE PROTEIN GENES, REI
? TITLE: SPANCENCE, ILSP06958
? CURRENT APPLICATION NUMBER: US/09/862,802
? PRIORITY FILING DATE: 2002-01-10
? PRIOR APPLICATION NUMBER: US-04/11,470
? PRIORITY FILING DATE: 1998-07-08
? NUMBER OF SEQ ID NOS: 11
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 4
? LENGTH: 316
? TYPE: PRT
? ORGANISM: Unknown
? FEATURE:
? OTHER INFORMATION: mammalian protein
US-09-862-802-4

```

[illegible]

```
QY 251 FSWKGLSDNDEGTQWQVDSPLSPSEQKYNSSGPN-- SGNEDCAFSSGSG-WN 103
      :||||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 236 YIKMKLST--PTAKMKVIVTIVATISQEN-WKLEQPTKMWGCHTSGDRCAPHPPTQW 292
      :||| 11:||||| 11:||||| 11:||||| 11:|||||
QY 304 DNRQDVINWICK 316
      :||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 293 IDVQVQRYHWICE 305
      :||| 11:||||| 11:||||| 11:||||| 11:|||||

RESULT 2
US-09-862-802-10
: Sequence 10, Application US/09862802
: Patent No. US20020165346A1
: GENERAL INFORMATION:
: APPLICANT: Schering-Plough Corporation
: TITLE OF INVENTION: MAMMALIAN MEMBRANE PROTEIN GENES: PEPTIDE FRAGMENTS
: FILE REFERENCE: SF0695B
: CURRENT APPLICATION NUMBER: US/09/862,802
: PRIOR FILING DATE: 2002-01-10
: PRIOR APPLICATION NUMBER: US 09/111,470
: PRIOR FILING DATE: 1998-07-08
: NUMBER OF SEQ. ID NOS: 11
: SOFTWARE: Patent version 3.1
: SEQ. ID NO: 10
: LENGTH: 273
: TYPE: PRT
: ORGANISM: monkey
: FEATURE:
: OTHER INFORMATION: mammalian protein
US-09-862-802-10

Query Match 18.08, Score 305, DB 9, Length 273,
Best local similarity 29.38, Freq. No. 3, gc-17,
Matches 89; Conservative 49; Mismatches 96; Indels 70; Gaps 15:

QY 20 HGAIVGLLTFMLAGVAILVQVSKVSSISQESQEDALTYQNTQKAAGLSLKS 79
      :||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 22 NGPLPIQST--ILIVYIVGVGQNSKRFQDVLTRTD---FSNT-----SN 63
      :||| 11:||||| 11:||||| 11:||||| 11:|||||
QY 80 KLOIYQELTQKAAGELPEKSKIOEIYQELTRKAAGELPFESKIOEIYQELTRKA 139
      :||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 64 TVARI-QALIS-----QTSIPE---TASIKAEVSGF--KQFRQAVHSPM----- 104
      :||| 11:||||| 11:||||| 11:||||| 11:|||||
QY 140 AVNTPPKSLQEIYQELTRKAAGVDELPRQSKQNTYQITDQKTAEPDLCPICTQW 199
      :||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 105 -----LKVQIVQDQKRLIQVAVTLNNKE-----ASTEGTC--CPVWV 143
      :||| 11:||||| 11:||||| 11:||||| 11:|||||
QY 200 FFGINCYEKSNSQPNMIDVTAGQFPAQIVTKTAEEQNPLOIQTSPNSPSSWQISQ 259
      :||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 144 EHQUSCTWFHSGMSNAEAKYQQLKNAHLVINSREKQNFYQKYL--GSATIMKGLSD- 200
      :||| 11:||||| 11:||||| 11:||||| 11:|||||
QY 260 NQGVITWQVQSPISPSQPYWNSCHPNN-----SGNEDCAFSSGSG-WNINPQVINY 312
      :||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 201 -PTAKMKVIVTIVATISQEN-WKLEQPTKMWGCHTSGDRCAPHPPTQW 292
      :||| 11:||||| 11:||||| 11:||||| 11:|||||
QY 413 WICK 316
      :||| 11:||||| 11:||||| 11:||||| 11:|||||
Db 259 WICE 262
      :||| 11:||||| 11:||||| 11:||||| 11:|||||

RESULT 3
US-09-978-295A-231
: Sequence 231, Application US/09978295A
: Patent No. 982002015606A1
: GENERAL INFORMATION:
: APPLICANT: Asitkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Bolstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleone
: APPLICANT: Filvaroli, Ellen
: APPLICANT: Fon, Sherman
```

```
: APPLICANT: Gao, Wei-Guang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gottschon, Mary E.
: APPLICANT: Hoddard, Audrey
: APPLICANT: Hoddowski, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kijavitt, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Rapley, Mary A.
: APPLICANT: Fan, James
: APPLICANT: Facol, Nicholas P.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Williams, Daniel
: APPLICANT: Williams, P. M. Key
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleo
: TITLE OF INVENTION: Acids Encoding the Same
: FILE REFERENCE: P26301C11
: CURRENT APPLICATION NUMBER: US/09/978,295A
: PRIOR FILING DATE: 2001-10-15
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-10
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066204
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077459
: PRIOR FILING DATE: 1998-04-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-04-11
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: PRIOR APPLICATION NUMBER: 60/078004
: PRIOR FILING DATE: 1998-03-13
: PRIOR APPLICATION NUMBER: 60/078886
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078936
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078910
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078939
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/079294
: PRIOR FILING DATE: 1998-03-25
: PRIOR APPLICATION NUMBER: 60/079656
: PRIOR FILING DATE: 1998-03-26
: PRIOR APPLICATION NUMBER: 60/079664
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079689
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079663
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079728
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079786
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079920
: PRIOR FILING DATE: 1998-03-30
: PRIOR APPLICATION NUMBER: 60/079923
: PRIOR FILING DATE: 1998-03-30
: PRIOR APPLICATION NUMBER: 60/080105
```


Tue Dec 10 10:10:09 2002

us-09-831-458a-12.std.rapb

Page 3

2 PRIOR FILING DATE: 1998-04-11
2 PRIOR APPLICATION NUMBER: 60/080107
2 PRIOR FILING DATE: 1998-03-31
2 PRIOR APPLICATION NUMBER: 60/080105
2 PRIOR FILING DATE: 1998-04-11
2 PRIOR APPLICATION NUMBER: 60/080194
2 PRIOR FILING DATE: 1998-03-31
2 PRIOR APPLICATION NUMBER: 60/080427
2 PRIOR FILING DATE: 1998-04-01
2 PRIOR APPLICATION NUMBER: 60/080328
2 PRIOR FILING DATE: 1998-04-01
2 PRIOR APPLICATION NUMBER: 60/080333
2 PRIOR FILING DATE: 1998-04-01
2 PRIOR APPLICATION NUMBER: 60/080334
2 PRIOR FILING DATE: 1998-04-01
2 PRIOR APPLICATION NUMBER: 60/081070
2 PRIOR FILING DATE: 1998-04-08
2 PRIOR APPLICATION NUMBER: 60/081049
2 PRIOR FILING DATE: 1998-04-08
2 PRIOR APPLICATION NUMBER: 60/081071
2 PRIOR FILING DATE: 1998-04-08
2 PRIOR APPLICATION NUMBER: 60/081195
2 PRIOR FILING DATE: 1998-04-08
2 PRIOR APPLICATION NUMBER: 60/081203
2 PRIOR FILING DATE: 1998-04-09
2 PRIOR APPLICATION NUMBER: 60/081229
2 PRIOR FILING DATE: 1998-04-09
2 PRIOR APPLICATION NUMBER: 60/081455
2 PRIOR FILING DATE: 1998-04-15
2 PRIOR APPLICATION NUMBER: 60/081817
2 PRIOR FILING DATE: 1998-04-15
2 PRIOR APPLICATION NUMBER: 60/081819
2 PRIOR FILING DATE: 1998-04-15
2 PRIOR APPLICATION NUMBER: 60/081952
2 PRIOR FILING DATE: 1998-04-15
2 PRIOR APPLICATION NUMBER: 60/081838
2 PRIOR FILING DATE: 1998-04-15
2 PRIOR APPLICATION NUMBER: 60/082568
2 PRIOR FILING DATE: 1998-04-21
2 PRIOR APPLICATION NUMBER: 60/082569
2 PRIOR FILING DATE: 1998-04-21
2 PRIOR APPLICATION NUMBER: 60/082704
2 PRIOR FILING DATE: 1998-04-22
2 PRIOR APPLICATION NUMBER: 60/082804
2 PRIOR FILING DATE: 1998-04-22
2 PRIOR APPLICATION NUMBER: 60/082700
2 PRIOR FILING DATE: 1998-04-22
2 PRIOR APPLICATION NUMBER: 60/082797
2 PRIOR FILING DATE: 1998-04-22
2 PRIOR APPLICATION NUMBER: 60/082796
2 PRIOR FILING DATE: 1998-04-23
2 PRIOR APPLICATION NUMBER: 60/083336
2 PRIOR FILING DATE: 1998-04-27
2 PRIOR APPLICATION NUMBER: 60/083322
2 PRIOR FILING DATE: 1998-04-28
2 PRIOR APPLICATION NUMBER: 60/083392
2 PRIOR FILING DATE: 1998-04-29
2 PRIOR APPLICATION NUMBER: 60/083445
2 PRIOR FILING DATE: 1998-04-29
2 PRIOR APPLICATION NUMBER: 60/083446
2 PRIOR FILING DATE: 1998-04-29
2 PRIOR APPLICATION NUMBER: 60/083499
2 PRIOR FILING DATE: 1998-04-29
2 PRIOR APPLICATION NUMBER: 60/083545
2 PRIOR FILING DATE: 1998-04-29
2 PRIOR APPLICATION NUMBER: 60/083554
2 PRIOR FILING DATE: 1998-04-29
2 PRIOR APPLICATION NUMBER: 60/083558
2 PRIOR FILING DATE: 1998-04-29
2 PRIOR APPLICATION NUMBER: 60/083559
2 PRIOR FILING DATE: 1998-04-29
2 PRIOR APPLICATION NUMBER: 60/083500
2 PRIOR FILING DATE: 1998-04-29

2 PRIOR APPLICATION NUMBER: 60/083742
2 PRIOR FILING DATE: 1998-04-30
2 PRIOR APPLICATION NUMBER: 60/084366
2 PRIOR FILING DATE: 1998-05-05
2 PRIOR APPLICATION NUMBER: 60/084414
2 PRIOR FILING DATE: 1998-05-06
2 PRIOR APPLICATION NUMBER: 60/084441
2 PRIOR FILING DATE: 1998-05-06
2 PRIOR APPLICATION NUMBER: 60/084637
2 PRIOR FILING DATE: 1998-05-07
2 PRIOR APPLICATION NUMBER: 60/084639
2 PRIOR FILING DATE: 1998-05-07
2 PRIOR APPLICATION NUMBER: 60/084640
2 PRIOR FILING DATE: 1998-05-07
2 PRIOR APPLICATION NUMBER: 60/084598
2 PRIOR FILING DATE: 1998-05-07
2 PRIOR APPLICATION NUMBER: 60/084600
2 PRIOR FILING DATE: 1998-05-07
2 PRIOR APPLICATION NUMBER: 60/084627
2 PRIOR FILING DATE: 1998-05-07
2 PRIOR APPLICATION NUMBER: 60/084643
2 PRIOR FILING DATE: 1998-05-07
2 PRIOR APPLICATION NUMBER: 60/085339
2 PRIOR FILING DATE: 1998-05-13
2 PRIOR APPLICATION NUMBER: 60/085348
2 PRIOR FILING DATE: 1998-05-13
2 PRIOR APPLICATION NUMBER: 60/085323
2 PRIOR FILING DATE: 1998-05-13
2 PRIOR APPLICATION NUMBER: 60/085582
2 PRIOR FILING DATE: 1998-05-15
2 PRIOR APPLICATION NUMBER: 60/085700
2 PRIOR FILING DATE: 1998-05-15
2 PRIOR APPLICATION NUMBER: 60/085689
2 PRIOR FILING DATE: 1998-05-15
2 PRIOR APPLICATION NUMBER: 60/085579
2 PRIOR FILING DATE: 1998-05-15
2 PRIOR APPLICATION NUMBER: 60/085580
2 PRIOR FILING DATE: 1998-05-15
2 PRIOR APPLICATION NUMBER: 60/085573
2 PRIOR FILING DATE: 1998-05-15
2 PRIOR APPLICATION NUMBER: 60/085704
2 PRIOR FILING DATE: 1998-05-15
2 PRIOR APPLICATION NUMBER: 60/085697

Query Match 17.1% Score 284.5, DB 9, Length
Best Local Similarity 29.4%, Prod No 7 1e-16,
Matches 91: Conserved 47: Mismatches 111, Sub 1

QY 23 LVLQSLSPMLASVIAIVIVVSVKVSSTISQVSRQDAIYQNTOLKAAV
DB 34 LALAVLTVTVMVILSL--TSKATFPVAILDGHDLKTNASKQITAI
QY 79 -SKLGEIYQELTOLKAAVGEIPEKSKLGEIYQELKAAVGEIPEKSKL
DB 92 -HSCSGTQAGLCIIPAEIGF--AAKIME-----QESAL
QY 138 KAAVGEIPEKSKLGEIYQELKAAVGEIPEKSKLGEIYQELKAAV
DB 134 LKVAAGKKE-----DKTELFRALEAV-RLONN-----
QY 198 WFFVQVNCYFMSNSQPMKHISVIAVQVVAQAVVITAFQNFICQNSP
DB 169 WLSFSCYFFSVFTTMAAGQVAVASNAIVIVSVLEVFITFN--
QY 258 DLNGESI--WLVVGSSTSPFGVYNSQEPNNS GNDKAEFSQGVV
DB 227 AVPHLQKVGIVGVIVVIVSVS--HWVGGFPMQWGRNVCWMLDITGLAV
QY 312 YWICKPPAC 321
DB 284 GWICKERHNC 293

RESULT 4
US-09-978-697-231
Sequence 231, Application US/09/7667
Patent No. US20020169284A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Bolstein, David
APPLICANT: Desnoyers, Ivar
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gertlson, Mary F.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavits, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James J.
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tomas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleotides
FILE REFERENCE: P26391027
CURRENT APPLICATION NUMBER: US/09/978 697
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 09/018085
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077652
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078436
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27

PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-29

```

PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083494
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083560
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084558
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697

```

```

Query Match 17.1% Score 289.5, PA 9, Length 293;
Best Local Similarity 29.4%; Pred No. 7 1e-16;
Matches 91; Conservative 47; Mismatches 11; Indels 61, Gaps 13;

```

```

QY 23 LVLQSLFMLAGVLAIVLVQSVSLSSQUSFVQVAVYDNLGLYKAAVTELSTK... 78
DB 34 LALAVLTIVIAVILITL--LSKASTERALLDDHDLRLNASKUTALGALKKEEYDGC 91
QY 79 -SKQEIYQELTOLKAAGVGLPEKSKQEIYQELTOLKAAGVGLPEKSKQEIYQELTOL 137
DB 92 HSCCSGVOALQOTTRALGCG--AQAKLME-----QESALELKLKSVTQCG 133
QY 138 KAAVGLPEKSKQEIYQELTOLKAAGVGLPEKSKQEIYQELTOLKAAGVGLPEKSKQEIY 197
DB 134 LAEGAGHGE-----DVRTLEFALAEAV-RIVQNNSS-----CEPCTTS 168

```

```

QY 198 WFFQCGCYFMSNSORNMHDSVLAQVFAVRAQIVVITKAFQONFLQIQTSSK...
DB 169 WLFESCSYFFSVFKPTWMAAGGHCATASALLVIVGSLDEGCFLENT--F
QY 258 DLNDEST...WQVDCSLSTSFQRYNSSEPNNS--GNFPAEFGSG--WV
DB 227 AVHLEKVGQYQWVDGVSLSST--HNNQCEYNDNAWPEHNCVMMIHTGIVL
QY 312 YWICKRPAAC 321
DB 284 GWICERHNC 293

RESULT 5
US 978 192A-231
Sequence 231, Application US/978/978192A
Patent No. US2002017753A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Bolstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gertelsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Gudowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Guiney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Pauli, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Thomas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William T.
TITLE OF INVENTION: Secreted and Transmembrane Polypept...
FILE REFERENCE: P2630PIC9
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: US/978/978192A
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886

```

PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078445
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078410
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079954
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079489
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079746
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081917
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081919
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081938
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082706
PRIOR FILING DATE: 1998-04-22

PRIOR APPLICATION NUMBER: 60/082747
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083346
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083544
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084543
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697

Query Match 17.18; Score 289.5; DB 9; Length 293;
Host Identity 29.48; Fred No. 7; le-16;
Matches 91; Conservative 47; Mismatches 111; Indels 61; Gaps 13;

```

OY 23 LVLLLSFMLAGVLAIVLVQVSKVPSLSQSEQUAIVNLIQLKAAGLESEK-----78
DB 34 LAAVAVTTVLAIVLITL-----LSKASTERAALLDGHDLRTNSKOTAAAGALKREYDGC 91
OY 79 -SKIAEIVLDELPAALAAVDELPERKSKIGETIVLDELKAAGVDELPERKSKIGETIVL 137
DB 92 HSNQSTIQAGLQITKAELVH-----AAKALME-----GNSALREIPEVYTG 133
OY 138 KAAGVDELPERKSKIGETIVLDELKAAGVDELPERKSKIGETIVLDELKAAGVDELPERKSKIGETIVL 197
DB 134 LAAGGKE-----DVRETELEFALEAV-RLOONS-----CEBPPPS 168
OY 198 WFFPUN-YFMSNSQFNMHNSVLAQVFAKLVVILKAEPANFLOLQSPNSFFSWGLS 257
DB 169 WLSFESQYFFSVYKTIKMAALH-ATASAHVIVVGLDQGLTETKNT-----KSGYV 226
OY 258 WLMOST---WQVWQSPILSPSPQYVWNSPENNNS-CHNECAFESDSSG-WNNPCTVFN 311
DB 227 AVPHLEKVGQYQWVDQVSLSPS---HMNQCFPNQWQFNFVMMIHTGIMNDAPQNSPKD 283
OY 312 YWICKKPAAC 321
DB 284 CMICEKKNHC 293

```

```

RESULT 6
US-09-864-761-40945
; Sequence 40945, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hancel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Assembly-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/753,595
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 68 24263.5
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/246,454
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00066
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00067
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00064
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00069
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00065
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00068
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00063
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00062
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00061
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00070
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/508,408

```

```

; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 40945
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC038812.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN ADIP T-LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HEPA1, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN PONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN TUNG, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: SWISSPROT HIT: P26717, EVALU6.6 G06-C
; OTHER INFORMATION: EST_HUMAN HIT: 60468.1, EVALU6.2.00
US-09-864-761-40945

```

```

Query Match 16.7% Score 284; DB 10; Length
Best Local Similarity 108.8%; Hitd. No. 2; 3e 16;
Matches 50; Conservative 0; Mismatches 0; Indels

```

```

OY 189 KLCHEPKIMTFEUNCYFMSNSQFNMHNSVLAQVFAKLVVILKAEPANFLOLQSPNSFFSWGLS
DB 1 KLCHEPKIMTFEUNCYFMSNSQFNMHNSVLAQVFAKLVVILKAEPANFLOLQSPNSFFSWGLS

```

```

RESULT 7
US-09-764-870-279
; Sequence 279, Application US/09764870
; Patent No. US2002004386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: NO-1-to-Antibis, Fibrinins, and Antibody
; FILE REFERENCE: PT214
; CURRENT APPLICATION NUMBER: US/09/764,870
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION DATA REMOVED - CONSULT PALM OF FILE 0799
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: Patonlin Ver. 2.0
; SEQ ID NO 279
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-870-279

```

```

Query Match 16.2% Score 274.5; DB 10; Length
Best Local Similarity 27.2%; Hitd. No. 1; 4e-14;
Matches 81; Conservative 51; Mismatches 123; Indels

```

```

OY 36 VLVAVLVVQVSKVPSLSQSEQUAIVNLIQLKAAGVDELPERKSKIGETIVL
DB 58 EVASVLLQVAVLVVQVSKVPSLSQSEQUAIVNLIQLKAAGVDELPERKSKIGETIVL
OY 96 GELP-EKSKIGETIVLDELKAAGVDELPERKSKIGETIVLDELKAAGVDELPERKSKIGETIVL
DB 112 VQVQVWNSQYVYKSPQIKVSV-----KRAKAVQVLLKSWEEVSLVNA
OY 150 LGFVYGETIKAAGVDELPERKSKIGETIVLDELKAAGVDELPERKSKIGETIVL
DB 158 LKASALNLIKALVLSLFLNMSLILPQNDITQVVS-----
OY 206 YFMSNSQFNMHNSVLAQVFAKLVVILKAEPANFLOLQSPNSFFSWGLS
DB 212 YFMSNSQFNMHNSVLAQVFAKLVVILKAEPANFLOLQSPNSFFSWGLS
OY 266 QWVWQSPILSPSPQYVWNSPENNNS-CHNECAFESDSSG-WNNPCTVFN
DB 276 SWVWQSPILSPSPQYVWNSPENNNS-CHNECAFESDSSG-WNNPCTVFN

```

RESULT 8

US-09-905-291A-377

Sequence 377, Application US/09905291A

Patent No. US20020160474A1

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.

APPLICANT: Ashkenazi, Avi

APPLICANT: Rotstein, David

APPLICANT: Desnoyers, Luc

APPLICANT: Patton, Dan L.

APPLICANT: Ferrara, Napoleone

APPLICANT: Filvaroff, Ellen

APPLICANT: Fong, Sherman

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerber, Hanspeter

APPLICANT: Gottschalk, Mary E.

APPLICANT: Godowski, Paul J.

APPLICANT: Grimaldi, Christopher J.

APPLICANT: Gunney, Austin L.

APPLICANT: Hillan, Kenneth, J.

APPLICANT: Kijavitt, Ivar J.

APPLICANT: Mather, Jennie P.

APPLICANT: Pan, James

APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann

APPLICANT: Stewart, Timothy A.

APPLICANT: Thomas, Daniel

APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William, I.

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

FILE REFERENCE: 10466-14

CURRENT APPLICATION NUMBER: US/09/905,291A

CURRENT FILING DATE: 2001-07-12

PRIOR APPLICATION NUMBER: PCT/US00/04414

PRIOR FILING DATE: 2000-02-22

PRIOR APPLICATION NUMBER: US 60/143,048

PRIOR FILING DATE: 1999-07-07

PRIOR APPLICATION NUMBER: US 60/145,698

PRIOR FILING DATE: 1999-07-26

PRIOR APPLICATION NUMBER: US 60/146,222

PRIOR FILING DATE: 1999-07-28

PRIOR APPLICATION NUMBER: PCT/US99/20594

PRIOR FILING DATE: 1999-09-08

PRIOR APPLICATION NUMBER: PCT/US99/20944

PRIOR FILING DATE: 1999-09-13

PRIOR APPLICATION NUMBER: PCT/US99/21090

PRIOR FILING DATE: 1999-09-15

PRIOR APPLICATION NUMBER: PCT/US99/21547

PRIOR FILING DATE: 1999-09-15

PRIOR APPLICATION NUMBER: PCT/US99/22089

PRIOR FILING DATE: 1999-10-05

PRIOR APPLICATION NUMBER: PCT/US99/28014

PRIOR FILING DATE: 1999-11-29

PRIOR APPLICATION NUMBER: PCT/US99/28313

PRIOR FILING DATE: 1999-11-30

PRIOR APPLICATION NUMBER: PCT/US99/28564

PRIOR FILING DATE: 1999-12-02

PRIOR APPLICATION NUMBER: PCT/US99/28565

PRIOR FILING DATE: 1999-12-02

PRIOR APPLICATION NUMBER: PCT/US99/40066

PRIOR FILING DATE: 1999-12-16

PRIOR APPLICATION NUMBER: PCT/US99/40911

PRIOR FILING DATE: 1999-12-20

PRIOR APPLICATION NUMBER: PCT/US99/40999

PRIOR FILING DATE: 1999-12-20

PRIOR APPLICATION NUMBER: PCT/US99/40916

PRIOR FILING DATE: 2000-01-05

NUMBER OF SEQ ID NOS: 423

SEQ ID NO: 377

LENGTH: 219
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-09-905-291A-377

Query Match:

16.1% Score 274.5; DB 9; Length 219;

Best Local Similarity 38.5% Prod. No. 9.1e-15;

Matches 60; Conserved 25; Mismatches 56; Indels 15; Gaps 6;

QY 172 KGGGVVAPLITLH-----IAEPLVPHVYKQWYTHQGNVYFMSNSQPMWHSVYVQVYV 276
 DB 55 KKEPLFENFETLSYNGSGSVKNC--CPINWEYFSSCYFSLDTLSWALSLKNCAMG 112
 QY 227 AGIVVKTAFEGHITQIQISQSNQPMWHSVYVQVYVYVYVYVYVYVYVYVYVYVYV 286
 DB 113 AHVAVINSQFFQFPLSYKPKKMPF-FVGLSQVYVQVYVYVYVYVYVYVYVYVYVYVYV 170
 QY 287 NNSGN-ELCAEFSLS-----GMDNMDVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 316
 DB 171 NNIALEDCATMKRSSNPQNMNDVTCPLNYPRICE 206

RESULT 9

US-09-909-320-377

Sequence 377, Application US/09909320

Patent No. US20020132240A1

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.

APPLICANT: Ashkenazi, Avi

APPLICANT: Rotstein, David

APPLICANT: Desnoyers, Luc

APPLICANT: Patton, Dan L.

APPLICANT: Ferrara, Napoleone

APPLICANT: Filvaroff, Ellen

APPLICANT: Fong, Sherman

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerber, Hanspeter

APPLICANT: Gottschalk, Mary E.

APPLICANT: Godowski, Paul J.

APPLICANT: Grimaldi, Christopher J.

APPLICANT: Gunney, Austin L.

APPLICANT: Hillan, Kenneth, J.

APPLICANT: Kijavitt, Ivar J.

APPLICANT: Mather, Jennie P.

APPLICANT: Pan, James

APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann

APPLICANT: Stewart, Timothy A.

APPLICANT: Thomas, Daniel

APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William, I.

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

FILE REFERENCE: 10466-14

CURRENT APPLICATION NUMBER: US 60/146,222

CURRENT FILING DATE: 2002-01-04

PRIOR APPLICATION NUMBER: PCT/US00/04414

PRIOR FILING DATE: 2000-02-22

PRIOR APPLICATION NUMBER: US 60/143,048

PRIOR FILING DATE: 1999-07-07

PRIOR APPLICATION NUMBER: US 60/145,698

PRIOR FILING DATE: 1999-07-26

PRIOR APPLICATION NUMBER: PCT/US99/20594

PRIOR FILING DATE: 1999-07-28

PRIOR APPLICATION NUMBER: PCT/US99/20944

PRIOR FILING DATE: 1999-09-08

PRIOR APPLICATION NUMBER: PCT/US99/20944

PRIOR FILING DATE: 1999-09-13

PRIOR APPLICATION NUMBER: PCT/US99/21090

PRIOR FILING DATE: 1999-09-15

PRIOR APPLICATION NUMBER: PCT/US99/21547

PRIOR FILING DATE: 1999-09-15

APPLICANT: Pan,James
APPLICANT: Smith,Victoria
APPLICANT: Matarabe,Colin K.
APPLICANT: Wood,William L.
APPLICANT: Zhang,Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEOTIDE
FILE REFERENCE: P3430101
CURRENT APPLICATION NUMBER: 08/10/052,596
CURRENT FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062450
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/064121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/064540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/064541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063564
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063734
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063870
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066120
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066466
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066772
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069335
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069425
PRIOR FILING DATE: 1997-12-12
PRIOR APPLICATION NUMBER: 60/069870
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/068017
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078839
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333

PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/083222
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084039
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086023
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086486
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087098
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087298
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05


```

? PRIOR APPLICATION NUMBER: 60/088202
? PRIOR FILING DATE: 1998-06-05
? PRIOR APPLICATION NUMBER: 60/088212
? PRIOR FILING DATE: 1998-06-05
? PRIOR APPLICATION NUMBER: 60/088217
? PRIOR FILING DATE: 1998-06-05
? PRIOR APPLICATION NUMBER: 60/088326
? PRIOR FILING DATE: 1998-06-04
? PRIOR APPLICATION NUMBER: 60/088555
? PRIOR FILING DATE: 1998-06-09
? PRIOR APPLICATION NUMBER: 60/088722
? PRIOR FILING DATE: 1998-06-10
? PRIOR APPLICATION NUMBER: 60/088738
? PRIOR FILING DATE: 1998-06-10
? PRIOR APPLICATION NUMBER: 60/088743
? PRIOR FILING DATE: 1998-06-10
? PRIOR APPLICATION NUMBER: 60/088811
? PRIOR FILING DATE: 1998-06-10
? PRIOR APPLICATION NUMBER: 60/088824
? PRIOR FILING DATE: 1998-06-10
? PRIOR APPLICATION NUMBER: 60/088825
? PRIOR FILING DATE: 1998-06-10
? PRIOR APPLICATION NUMBER: 60/088826
? PRIOR FILING DATE: 1998-06-10
? PRIOR APPLICATION NUMBER: 60/088861
? PRIOR FILING DATE: 1998-06-11
? PRIOR APPLICATION NUMBER: 60/088863
? PRIOR FILING DATE: 1998-06-11
? PRIOR APPLICATION NUMBER: 60/088876
? PRIOR FILING DATE: 1998-06-11
? PRIOR APPLICATION NUMBER: 60/089090
? PRIOR FILING DATE: 1998-06-12
? PRIOR APPLICATION NUMBER: 60/089105
? PRIOR FILING DATE: 1998-06-12
? PRIOR APPLICATION NUMBER: 60/089512
? PRIOR FILING DATE: 1998-06-16
? PRIOR APPLICATION NUMBER: 60/089514
? PRIOR FILING DATE: 1998-06-16
? PRIOR APPLICATION NUMBER: 60/089538
? PRIOR FILING DATE: 1998-06-17
? PRIOR APPLICATION NUMBER: 60/089598
? PRIOR FILING DATE: 1998-06-17
? PRIOR APPLICATION NUMBER: 60/089653
? PRIOR FILING DATE: 1998-06-17
? PRIOR APPLICATION NUMBER: 60/089690
? PRIOR APPLICATION NUMBER: 60/089690

Query Match 16.1% Score 273.5; DB: 12; Length 219;
Best Local Similarity 38.5%; Freq. No. 9; Indels 15; Gaps 6;
Matches 60; Conservative 25; Mismatches 56; Indels 15; Gaps 6;

```

```

QY 172 KQOQVYQELTTLK...TAFERLPHKPIW:FGUNYHNSNSGNNMH:SVTA'GEVY 27%
DB 55 KKFQLEPMTTELSCYNVSGSVKNC--CPLMWEYFQSSCYFFSDTISMLSLKNS:AMG 112
QY 227 AOLVYIKAEQNF:QLOTSRNSRPSWMSJDLNQBGTWQVWDSPI SPSPQRYWNSGEP 286
DB 113 AHLVYINSQEDQFLSYKKPKMREF-FIGLSDYVEJQWQWVDSTPLTKSIS-FWDVGEP 170
QY 287 NNSGN-EDCAEFGS-----GNNDNRCDDVDNYWICK 316
DB 171 NNIALIHOCATIRROSSNRKNNNDYTCPLNFRICE 206

RESULT 12
US-09-862-802-5
? Sequence 5, Application US/09862802
? Patent No. US20020165346A1
? GENERAL INFORMATION:
? APPLICANT: Schering-Plough Corporation
? TITLE OF INVENTION: MAMMALIAN MEMBRANE PROTEIN GENES; RELATED REAGENTS
? FILE REFERENCE: SP0695B
? CURRENT APPLICATION NUMBER: US/09/862.802
? CURRENT FILING DATE: 2002-01-10

```

```

? PRIOR APPLICATION NUMBER: US 09/111,470
? PRIOR FILING DATE: 1998-07-04
? NUMBER OF SEQ ID NOS: 11
? SOFTWARE: Patent in version 4.1
? SEQ ID NO 5
? LENGTH: 291
? TYPE: PRT
? ORGANISM: Unknown
? FEATURE:
? OTHER INFORMATION: mammalian protein
US-09-862-802-5

```

```

Query Match 16.0% Score 271.5; DB: 9; Length
Best Local Similarity 27.0%; Freq. No. 1; Indels
Matches 86; Conservative 45; Mismatches 112; Indels
QY 6 EPRVQGLIGLGGALVGLVQLSTFMILAVIVALLVQVSVPSLSGEGS
DB 29 QPDLQRL-----CSGRLLILSL-----GLSLILVWVIVIGSUNSLQI
QY 66 TULKAAGVGLSEKSKIQETVYQETIQLKAAGVPIPKSKLQETVYQETILK
DB 78 SMTFAST-----FAVVKELSTGNNVGRKMSLE-----
QY 126 KLOEIVQELTLKAAGVGLPEKSKLOEIVQETIQLKAAGVGLPQSKQGG
DB 108 QLEKQKDLSE-----DHSLSL-----LLHVKQFVSPDLRSLSIQMAV
QY 186 ATERLCRPHCKPWTFEGGNCYHNSNSGNNMH:SVTA'GEVYQVLRJY
DB 148 GSERIC--CPVWNVVHEHSCYWPFSYKAMADANVCLDLAHLVYVTSW
QY 246 SPSNPSWMSJDLNQBGTWQVWDSPI SPSPQRYWNSGEPNN-----S
DB 206 GFVN--TWMSLHD--QNGPWKWDGDIYETSPKN-WPEQDDWYGHJL
QY 300 SG-WINDNRCDDVDNYWICK 316
DB 261 DGRWDDVQRPYRWCE 278

```

```

RESULT 13
US-09-862-802-6
? Sequence 6, Application US/09862802
? Patent No. US20020165346A1
? GENERAL INFORMATION:
? APPLICANT: Schering-Plough Corporation
? TITLE OF INVENTION: MAMMALIAN MEMBRANE PROTEIN GENES; REI
? FILE REFERENCE: SP0695B
? CURRENT APPLICATION NUMBER: US/09/862.802
? CURRENT FILING DATE: 2002-01-10
? PRIOR APPLICATION NUMBER: US 09/111,470
? PRIOR FILING DATE: 1998-07-04
? NUMBER OF SEQ ID NOS: 11
? SOFTWARE: Patent in version 4.1
? SEQ ID NO 6
? LENGTH: 287
? TYPE: PRT
? ORGANISM: Unknown
? FEATURE:
? OTHER INFORMATION: mammalian protein
US-09-862-802-6

```

```

Query Match 15.8% Score 267.5; DB: 9; Length
Best Local Similarity 28.8%; Freq. No. 3; Indels
Matches 92; Conservative 45; Mismatches 103; Indels
QY 6 EPRVQGLIGLGGALVGLVQLSTFMILAVIVALLVQVSVPSLSGEGS
DB 28 QPDLQRLSMWQFS-----LALSHRL-----LLVIVTIGSQAQV
QY 66 TULKAAGVGLSEKSKLQETVYQETIQLKAAGVPIPKSKLQETVYQETILK

```

```

db 70 KSLKEAFNS-SSTLEFV-QALSTHGSGVD-----KITSIGA----- 106
QY 126 KLQELVQHLRLKAAGELPRKSLQELIYQHLRLKAAGELVUSKQVQVYQHLIHLKI 185
db 107 KLEQOOD--LKAADHDL-----LPELKHFPVDLFEVACQWELH-----SN 145
QY 186 APERLCHCHCKEDVFFPGNGYFMSNSQPMHNSVTACQEVPAQLVVIKIAEONFLQGLQ 245
db 147 GSKRTG--CPVNWVERHQSCTYWSHSHSKAMAPKYGQQLFNAILVVTNSMTYQKTVQHT 204
QY 246 SPSNPFESMGLSDLNQELTWQWVQSPFISYFQYVNSDEFNN-----SCNEDCAEPG 209
db 205 NPEN--TWIGITD--SDGSWKVWDGTDYRHNRYN-WAVTQPDQWGHSHPLGSGEDVEVAF 259
QY 300 SG-WNDNRCDVDNYWICK 317
db 260 TCFNRHRCQQLQVYRWVCFK 278

```

```

RESULT 14
US-10-090-466-4
: Sequence 4: Affiliation US/10090466
: Patent No. US20020137914A1
: GENERAL INFORMATION:
: APPLICANT: Turner, C. Alexander Jr.
: APPLICANT: Matbur, Brian
: APPLICANT: Gullinan, Emily B.
: TITLE OF INVENTION: No. US20020137914A1-1 Human Protein and Polynucleotide
: FILE REFERENCE: LEX-0315-USA
: CURRENT APPLICATION NUMBER: US/10/090,466
: PRIOR FILING DATE: 2002-03-01
: PRIOR APPLICATION NUMBER: US 60/274,961
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO: 4
: LENGTH: 182
: TYPE: PRT
: ORGANISM: homo sapiens
: FEATURE:
: NAME/KEY: VARIANT
: LOCATION: (1)...(182)
: OTHER INFORMATION: Xaa - Any Amino Acid
US-10-090-466-4

```

```

Query Match: 15.5%; Score 263.5; DB 12; Length 182;
Best Local Similarity 35.9%; Pred. No. 4,4e-14;
Matches 65; Conservative 27; Mismatches 66; Indels 24; Gaps 8;

```

```

QY 144 LPEKSLQELIYQELTRKAAGVETPIYQSKQVQVYQHT--DLKTAPEPLCPHCKWMTF 200
db 11 VPHNPMYSKIVKRLSKLR-----EYQYHXSILTCVMECKDTEHWSG--CPTPWTS 58
QY 201 PGNVYFMSNSQPMHNSVTACQEVPAQLVVIKIAEONFLQGLQSPS-NPFSMMGLSDLN 260
db 59 FQSSCYFTISGMWSKSKQNMVMAADLVVITNREDOFT-IQNLKNSSYFLGLSDPG 117
QY 261 QGTWQWVQSPFISYFQYVNSDEFNN--EFGSG--GWNDRCDVDNYWIC 315
db 118 GRHMQWVQDTPYNFNV-TFMHSGEPNNL-DEFCALINPSSFEWQWMDIHCHVQKSLG 175
QY 316 K 316
db 176 K 176

```

```

RESULT 15
US-10-090-466-2
: Sequence 2: Affiliation US/10090466
: Patent No. US20020137914A1
: GENERAL INFORMATION:
: APPLICANT: Turner, C. Alexander Jr.
: APPLICANT: Matbur, Brian

```

```

: APPLICANT: Gullinan, Emily B.
: TITLE OF INVENTION: No. US20020137914A1-1 Human Protein and Polynucleotide
: FILE REFERENCE: LEX-0315-USA
: CURRENT APPLICATION NUMBER: US/10/090,466
: PRIOR FILING DATE: 2002-03-01
: PRIOR APPLICATION NUMBER: US 60/274,961
: PRIOR FILING DATE: 2001-03-12
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO: 2
: LENGTH: 213
: TYPE: PRT
: ORGANISM: homo sapiens
US-10-090-466-2

```

```

Query Match: 15.5%; Score 263.5; DB 12; Length 213;
Best Local Similarity 35.9%; Pred. No. 5,4e-14;
Matches 65; Conservative 27; Mismatches 66; Indels 23; Gaps 8;

```

```

QY 144 LPEKSLQELIYQELTRKAAGVETPIYQSKQVQVYQHT--DLKTAPEPLCPHCKWMTF 200
db 42 VPHNPMYSKIVKRLSKLR-----EYQYHXSILTCVMECKDTEHWSG--CPTPWTS 89
QY 201 PGNVYFMSNSQPMHNSVTACQEVPAQLVVIKIAEONFLQGLQSPS-NPFSMMGLSDLN 260
db 59 FQSSCYFTISGMWSKSKQNMVMAADLVVITNREDOFT-IQNLKNSSYFLGLSDPG 148
QY 261 QGTWQWVQSPFISYFQYVNSDEFNN--EFGSG--GWNDRCDVDNYWIC 315
db 118 GRHMQWVQDTPYNFNV-TFMHSGEPNNL-DEFCALINPSSFEWQWMDIHCHVQKSLG 206
QY 316 K 316
db 207 K 207

```

```

Search completed: December 7, 2002, 11:08:16
Job time : 24 secs

```



```

CURRENT APPLICATION DATA:
: APPLICATION NUMBER: 09/08/365,103B
: FILING DATE: 28-DEC-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Nebel, Heidi S.
: REGISTRATION NUMBER: 37,719
: REFERENCE/DOCKET NUMBER: 0111 N5-24
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (515) 288-4667
: TELEFAX: (515) 288-1348
: INFORMATION FOR SEQ. ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 300 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-365-103B-6

```

```

Query Match: 19.94; Score: 338; DP: 1; Length: 300;
Best Local Similarity: 30.49; Prod. No. 2,10-23;
Matches: 93; Conservative: 59; Mismatches: 104; Indels: 50; Gaps: 14.

```

```

QY 34 AGVLAIVLYOVSKVPSLSQESQDAIYQNTLY... KAAVGELSEKSLQIYQEL 88
DB 11 AGLLALLIMHMETEKMLKQ---LGDIAIONVSHIVTKDLQPSNQLAKSQVYVMSQNL 67
QY 89 TOLKAAGVELPEKSKLOIYQITPLKAAGVELPEKSKLOIYQELTLK-----AAVG 142
DB 68 QELAEQKQK--KQDPSLSNLTGLQEDLPN--AQSNKSLSONLNLDLQDLVNIKSLG 123
QY 143 ELPEK---SKLOIYQELTLKAAGVELPDQSKQDIYQELDLTAIFERLRHFKPM 198
DB 124 LNKRIASLSLEKLEQHVAKLKEI...-----LISKQTA-----CNLEPKMW 163
QY 169 TPEQSNQYFMSNSQPMNHSVTAQVEPAQIVYIKTAEDPNTGLCTSPNPFSSNCTSF 258
DB 164 LHFQOKIYVYFGKSGKQWQIARFACSDLOGPLVSHSUKEDQDFLMQIHKKD--SWGLDD 221
QY 259 INQGTQWQVNSPLSPSPYWNSSPPNNSG--NPPAFPSGSG--WINDPQ--IVNRYWIC 315
DB 222 LIMELEFVMSRSPVYS---NMNPPFPNNPQCTET--VMNPISQWNTAFQPSVYIYAWVG 278
QY 316 KKPAC 321
DB 279 HOLATC 284

```

```

RESULT 5
: US-08-365-103B-4
: Sequence 4, Application US/08365103B
: Patent No. 5766943
: GENERAL INFORMATION:
: APPLICANT: Lynch, Richard G.
: APPLICANT: Nunez, Raphael D.
: APPLICANT: Yodanis, Jungi
: TITLE OF INVENTION: DNA Sequences for Soluble Forms of C123
: TITLE OF INVENTION: and Methods of Use for Same
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Zarley, McKee, Thome, Voorhees & Sease
: STREET: 801 Grand Ave. Suite 3200
: CITY: Des Moines
: STATE: Iowa
: COUNTRY: United States
: ZIP: 50309
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: 09/08/365,103B

```

```

FILING DATE: 28-DEC-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Nebel, Heidi S.
: REGISTRATION NUMBER: 37,719
: REFERENCE/DOCKET NUMBER: 0111 N5-24
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (515) 288-4667
: TELEFAX: (515) 288-1348
: INFORMATION FOR SEQ. ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 287 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-365-103B-4

```

```

Query Match: 19.74; Score: 344; DP: 1; Length:
Best Local Similarity: 31.96; Prod. No. 4,60-23;
Matches: 91; Conservative: 47; Mismatches: 101; Indels: 14;

```

```

QY 44 VSKVPSLSQESQDAIYQNTLY... KAAVGELSEKSLQIYQELTLK 88
DB 26 VSHVTKDLQPSNQLAKSQVYVMSQNL 67
QY 104 LQIYQITPLKAAGVELPEKSKLOIYQITPLKAAGVELPEKSKLOIYQELTLK 142
DB 75 LTIQDLEMAQSNKSLQ--NLRNLDLQDLVNIKSL--GLNKRKASDNT 123
QY 160 LKAAGVELPEKSKLOIYQITPLKAAGVELPEKSKLOIYQELTLK 198
DB 131 LMEI-----LISKQTA-----CNLEPKMW 163
QY 220 TACEVPAQIVYIKTAEDPNTGLCTSPNPFSSNCTSF 258
DB 170 FAFVTEVLSHSKQELLMHKKD--SWGLDD 221
QY 280 YWNSGPPNNSG--NPPAFPSGSG--WINDPQ--IVNRYWIC 315
DB 227 NMNPPFPNNPQCTET--VMNPISQWNTAFQPSVYIYAWVG 278

```

```

RESULT 6
: US-09-535-521-2
: Sequence 2, Application US/09535521
: Patent No. 6410714
: GENERAL INFORMATION:
: APPLICANT: Weber, Eric R.
: APPLICANT: McCall, Catherine A.
: TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IGE RECEPTOR
: TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES AND
: FILE REFERENCE: AL-5
: CURRENT APPLICATION NUMBER: 09/095355, 521
: CURRENT FILING DATE: 2000-03-24
: EARLIER APPLICATION NUMBER: 60/125,913
: EARLIER FILING DATE: 1999-04-24
: NUMBER OF SEQ. ID NOS: 26
: SOFTWARE: Patent In Ver. 2.1
: SEQ. ID NO: 2
: LENGTH: 292
: TYPE: PRT
: ORGANISM: Canis familiaris
: US-09-535-521-2

```

```

Query Match: 19.48; Score: 329.5; DP: 1; Length:
Best Local Similarity: 29.54; Prod. No. 1,20-23;
Matches: 97; Conservative: 45; Mismatches: 102; Indels: 14;

```

```

QY 11 QLTGCTGGAIVIGLSPNIAVVAHVSVKPSLSQESQDAIYQNTLY... KAAV 88
DB 27 QLTGCTGGAIVIGLSPNIAVVAHVSVKPSLSQESQDAIYQNTLY... KAAV 88
QY 70 AAVRTSEKSLQIYQITPLKAAGVELPEKSKLOIYQITPLKAAGVELPEKSKLOIYQELTLK 142

```

```

Query Match 19.4% Score 329.5% DR 4.4 Length 2627
Host Local Similarity 29.5% Read No. 1,26-227
Matches 972 Conservative 491 Mismatches 1021 Models 81 Gaps 15

07 11 GAGTCTCGGAGATVLLSTLMLAGVLAIVQVSKVSSLSQESQFDAYQVNIQIRK 69
   |||||
   |||||
08 27 CTAAGC-----LVYIMWAGVLTITLIM-----HRETV QHRIKLLIV 62
   |||||
   |||||
09 70 AAVGELSEKSLQETVOELTQIRKAAGELPEPSKIQETVOELTRIKAAVGLPEKSKLQV 129
   |||||
   |||||
10 64 AAVGVNSVNSGNDIEPHNG-----DQMAQSGQAQVSGD-----MKE 98
   |||||
   |||||
11 130 LYEELTRIKAAVGLPEKSKLQETVOELTRIKAAAGELPSV-----KQGLQYQELTRIKAE 187
   |||||
   |||||
12 99 IQAQRKKRKAQDLS-----QNLALNSDILNNIKSSQSLNEPSTALSHLEKLEPEV 149
   |||||
   |||||
13 07 EPI-----CPHGRKNTFTFGNLYEMNSQPMHDSVIAQVLEVAQVWIKTAE 256
   |||||
   |||||
14 150 FRIWMLHVSNGSRGNCQPEKKI NPEKRIYGFYGFPEPKKMIQARPAQSKLQGLQAS HNSQF 309
   |||||
   |||||
15 05 237 EQNLQQLQTSRNSNPFSSMGLSTEN NQPGYQAWVNGSPHSQEQFVNSQTEPANNQ-NETVA 295
   |||||
   |||||
16 210 EQDELAVYANKKQ--TWIETRI PDI DPEPFEFTIMENDPLANS--MMHDPEDPNQDQPEV 264
   |||||
   |||||
17 03 296 FPEPSG--WINDPQ--DIVDNYWIRKKVAAAG 321
   |||||
   |||||
18 265 MMAPSTGMAVNAFQSSLDG--WVVDPLATV 292

```

```

Sequence # Application US/08 465103B
Patent No. 576994
GENERAL INFORMATION:
APPLICANT: Lynch, Richard G.
APPLICANT: Nunez, Raphael D.
TITLE OF INVENTION: DNA Sequences for Soluble Forms of CD23
TITLE OF INVENTION, and Methods of Use for Same
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESS: 7411-y, McKee, Thermo, Voorhees & Sease
STREET: 861 Grand Ave., Suite 1200
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: ds/nd/365.103H
FILING DATE: 28-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Nabel, Heidi S.
REGISTRATION NUMBER: 37,719
PREFERENCE/WORKET NUMBER: 0115 24
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 288-3667
TELEFAX: (515) 288-1138
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-365-103B-8
Query Match: 19.0%, Score 322.5, DB 1, Length 321;
Best Local Similarity 28.6%; Pred. No. 6,36-22;
Matches 95, Conservative 61, Mismatches 121, Indels 55, Gaps 13;
Db 1 MSISRRFFVQGLLILDLITGHA--LVYQLSRLAGLYALVLVNSKRVSSLSQ-EQSE 57
1 MEHQVS-EIETPRRKRREGGIVITDVLTAEWAGLTLLLMHTTQSLEKOLEIRA 60
58 JAVYGH-QEKAWE-LISSEGTIVLHVEGLQKAAVDHLPRSKJ-DEYDELTLGA 116
61 APNVYSKRLPIHRIDQMAKSQTSGTSQEELERA-----EQPLRS 104
117 AVCHLPKKKELCFTYEPTTRFAAWEFLPRSK---IQELYQFTIRKAAV-TLTPQSK 172
105 Q--IHLSNNMLNIGADINSKRSQ--FINPVASDILFPI RFEVTYR----- 149
173 QQQVYGVTHYKAFAPLPCVPHIKMTWTFYTCYCMYMSNGNMHSNTVNOVEPAVYT 232
150 ---KELLVSSQVVRHCTELWMIR--PEVYVCKGIQWHAATV-LQEEGLVLS 201
223 KTAFTPTPLGEGSPSPNPFWMT-SKVNAFTWQWWDSPSTPSQPPYWSQDET INGN 291
202 HSHEDLI-KIASHSIY-SWIGLENLELKPLFMWGDSHYVS- NMATGDTLRSGSD 256
252 ETVAEITSGSV-WGDNQTYE-VHWICIEELAN 321
DB 257 EKVYMGQSTENMDATIELSLAWWDRLATG 288

```

```

GENERAL INFORMATION:
APPLICANT: Valladeau, Jenny
APPLICANT: Ravel, Odile
APPLICANT: Bates, Elizabeth E.M.
APPLICANT: Ford, John
APPLICANT: Saeland, Sem
APPLICANT: Lebecque, Serge J.E.
TITLE OF INVENTION: Mammalian Membrane Protein Genes;
TITLE OF INVENTION: Related Reagents
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESS: DNAX Research Institute
STREET: 901 California Avenue
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.40
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/111,470
FILING DATE: 08-JUL-1994
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 60/053,080
FILING DATE: 09-JUL-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: SP0645
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)952-9146
TELEFAX: (650)496-1200
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 316 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-111-470-4

Query Match 18.8%; Score 319.5; DB 4; Length 316;
Best Local Similarity 28.8%; Pred. No. 1,26-21;
Matches 90; Conservative 53; Mismatches 103; Indels 67; Gaps 13;

QY 23 LVQLLSFMLAGLVAILVGVSKVSSLSQSESDAIYVNLQLDKAAVGLSEKSKLV 82
DB 41 LLLSLGLILLVILVGVGFSKFRDLVLRITD...FSNLSINIVA-----ELQ 88
QY 83 ELVQELTDLKAAVGLPEKSKLDELVELTLKAAVGLPEKSKLDELVELTLKAAV 142
DB 89 ALTSQSSSLIEFTASL--KAEVGEKDE--RQAGVSEL-----DEHTQKAHLG 133
QY 144 ELPE-----KSLQELIYQELTRKAAVSEL PPGSKQVQIYQELTLKTAEPPL 190
DB 134 HCPHPSVCVPHVSEMLRLVQVQIVQDLKLKLCQVATLNNNA-----STFEGT 179
QY 191 CHCKCKDMTFPGQNTYFMSNSLPRMHHISVLAQGVPAQVVIKIAFQNTLQLISPSNP 250
DB 180 C--CVNVNVEHSDSCYVSHSISMVAEAKKYQALKNAHVVINSFPELNFVQVYT--GSA 236
QY 251 FSWMGISLDLNGGTQWQWDGSLSPSFQRVYNNSSPEPN-----SGNDFAEISGSG-NN 303
DB 246 YVWMLNLS--PPGAKKVVVCTIVATCPQN--WKPPQPPHWWYTHGICVPTFAAHFHPTGPMN 292
QY 304 DNKCDVDNWK 316
DB 293 DDVCFRPHWCE 305

```

```

RESULT 10
US-08-365-103B-10
Sequence 10, Application US/08/51048
Patent No. 5765943
GENERAL INFORMATION:
APPLICANT: Lynch, Richard G.
APPLICANT: Nance, Raphael D.
APPLICANT: Yodanis, Jungi
TITLE OF INVENTION: DNA Sequences for Soluble Forms of
TITLE OF INVENTION: and Methods of Use for Same
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESS: Zaffey, McKee, Thumte, Voorhees & Sease
STREET: 801 Grand Ave, Suite 3200
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,103B
FILING DATE: 28-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Nebel, Heidi S.
REGISTRATION NUMBER: 37,719
REFERENCE/DOCKET NUMBER: 01RT N5-24
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 288-9667
TELEFAX: (515) 288-1348
INFORMATION FOR SEQ ID NO: 103:
SEQUENCE CHARACTERISTICS:
LENGTH: 320 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-365-103B-10

Query Match 18.7%; Score 317.5; DB 1; Length
Best Local Similarity 29.0%; Pred. No. 1,86-21;
Matches 94; Conservative 59; Mismatches 116; Indels
QY 9 VQVLGLGLGHCAG--LVQLLSFMLAGLVAILVGVSKVSSLSQ--ELQ 82
DB 8 IELPRKRCQKQKQJYIVLGLVIALTWAGLLILLIMWHTQSLKQLEH 88
QY 66 TDLKAAVGLSEKSKLDELVELTLKAAVGLPEKSKLDELVELTLKAAV 142
DB 68 KNLSEHGDQMAKQKOSTOISQLEELRA-----FQSGT 133
QY 125 SKIQLIYQELTRKAAVGLPEKSKLDELVELTLKAAVGLPEKSKLDELVEL 190
DB 110 WNLNGVGLSEKSKQ--ELNEPNEASDLDELVELTREVTKLR----- 179
QY 181 TLKTAEPPLCPHCPKRWTFPGQNTYFMSNSLPRMHHISVLAQGVPAQVVIK 250
DB 149 MFLVNSQVWVNNIPELWVNVVPHVSHIENQTEFWVHAPVAGPDMVQVY 236
QY 241 LSLQPSRPFQSWMGISLDLNGGTQWQWDGSLSPSFQRVYNNSSPEPN----- 303
DB 209 LTKHSHHS--SWIGLRNLKLKCEFTIWDGSHVDYS--NNARPEPTSKG 292
QY 300 SG--WNNPQDVD--NPKCKKPAAG 321
DB 264 SGWMDAFCDKRLKAGVAGCDRLALG 287

```

RESULT 11
US-09-111-470-10

Sequence 10, Application US/09111470
Patent No. 6277959
GENERAL INFORMATION:
APPLICANT: Valladeau, Jenny
APPLICANT: Kavel, Odile
APPLICANT: Hates, Elizabeth E.M.
APPLICANT: Ford, John
APPLICANT: Saeland, Sem
APPLICANT: Lebecque, Serge J P
TITLE OF INVENTION: Mammalian Membrane Protein Genes,
TITLE OF INVENTION: Related Reagents
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSER: DNAX Research Institute
STREET: 901 California Avenue
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/111,470
FILING DATE: 08-JUL-1998
CLASSIFICATION:
Prior Application DATA:
APPLICATION NUMBER: US 60/053,080
FILING DATE: 09-JUL-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: SF0695
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)852-9196
TELEFAX: (650)496-1200
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 273 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-111-470-10

Query Match: 18.08; Score 305; DB 4; Length 273;
Best Local Similarity: 29.38; Pred No 2, Ids 20;
Matches: 89; Conservative 49; Mismatches 96; Indels 70; Gaps 15;

QY 20 HGAIVLQISPMITAGVIVALLVWVSKVPSISQSESDDAIYQNTOLKAVGELSEK 79
DB 22 NCLPLOSLL--LLVLIICVCGFQNSKRFQRIYLRID---FSNIT .. SN 63
QY 80 KLOFVYQELTOLKAVGELPEKSKLOFTYQELTRKAVGELPEKSKLOFTYQELTRKA 139
DB 64 TVAEI-QALIS-----QGSLEE--TASLKAEEGF--KQKRAVHEML----- 104
QY 140 AVGELPEKSKLOFTYQELTRKAVGELPEKSKQOYQELTDLKAFRELPHCKPDMT 199
DB 105 -----FVQGVVQIKKLTQVATLNNNGE-----ASTETG--CIVNVV 143
QY 200 FFGNCYTFMNSQRNHDSTVACQVEVAOLVKTAEQNFLOLOTSRNSRSGSLSD 259
DB 144 EHQNSYWFHSGMSMAEAKYVQIKNAHLVINSREQNFVQKYL--GSATVMGLSD- 200
QY 260 NQPTFMQVNDSPISPFQRYTWNSEPN-----SGNECAFESSG--WNIWPCGVNV 312
DB 201 -PFGAKKWDGTDIYATGPN-WKPGQPDNGYCHGSGGEPFAHFHPD-RRW-IVW-QFTYH 278
QY 413 WICK 316
DB 259 WICE 262

RESULT 12
US-08-688-442-4
Sequence 4, Application US/08688442
Patent No. 5821964
GENERAL INFORMATION:
APPLICANT: Au-Yang, Janice
APPLICANT: Cocks, Benjamin G.
APPLICANT: Goff, Surya K.
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: NOVEL HUMAN C-TYPE LECTIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSER: Jocyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/688,442
FILING DATE: Filed Herewith
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0045-1 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 292 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 1245724
US-08-688-442-4

Query Match: 17.78; Score 303.5; DB 2; Length 292;
Best Local Similarity: 28.68; Pred No 3, Ids 20;
Matches: 90; Conservative 54; Mismatches 98; Indels 73; Gaps 16;

QY 26 HGAIVLQISPMITAGVIVALLVWVSKVPSISQSESDDAIYQNTOLKAVGELSEK 68
DB 22 NCLPLOSLL--LLVLIICVCGFQNSKRFQRIYLRID---FSNIT .. SN 63
QY 80 KLOFVYQELTOLKAVGELPEKSKLOFTYQELTRKAVGELPEKSKLOFTYQELTRKA 138
DB 73 KLPSPFISNI-VAEI-QALIS-----QGSLEE--TASLKAEEGF--KQKQ 116
QY 139 EYQETTFPAAVGELPEKSKLOFTYQELTRKAVGELPEKSKQOYQELTDLKAFREL 188
DB 117 AVHEML-----LRVQIVQDIKKLTQVATLNNNGE-----ASTE 153
QY 199 PLPHCTFEWTFQGNVFMNSQSEWDSVTAQFVPAQIVVTFAEQNFLOLOTSEK 248
DB 154 GID--CFVNVVHFQDSYWFHSGMSMAEAKYVQIKNAHLVINSREQNFVQKYL--G 209
QY 249 NQPTFMQVNDSPISPFQRYTWNSEPN-----SGNECAFESSG-- 401
DB 210 SAYTVMGLSD--PEGAKKWDGTDIYATGPN-WKPGQPDNGYCHGSGGEPFAHFHPD 266
QY 302 WINDKQGVNDYWICK 316

Db 267 WNDVCOBPHWCE 281

RESULT 13

US-09-113-788-4
Sequence 4: Application US/09113788

Patent No. 5966104

GENERAL INFORMATION:

APPLICANT: Au-Young, Janice

APPLICANT: Cooks, Benjamin G

APPLICANT: Goll, Surya K.

APPLICANT: Hillman, Jennifer L.

TITLE OF INVENTION: NOVEL HUMAN C-TYPE LECTIN

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: US

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSO Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/113,788

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/668,342

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0095-1 CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-845-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 292 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 123724

US-09-113-788-4

Query Match 17.9% Score 303.5; DB 2: Length 292.

Best Local Similarity 28.6%; Pred. No. 3.1e-20;

Matches 90; Conservative 54; Mismatches 98; Indels 73; Gaps 16;

20 HCAVLQDL-----STMLA---GVLAIVGVSKVSSLSQESQDAIYQTLVL 68

Db 22 NCPILQSLQRLRSRCHLITSLGIIIVICVWCPQNSKQ-----PDLATL 72

QY 69 KAVVGLSEKSKLOEYQELTQKAAVGLPEKSKLOEYQELTRIKAAVGLPEKSKLTQ 128

Db 73 KTFDSNFTSNIT-VAEI-QALTS-----QSSSLF---TASIKAVESEF--KQERQ 116

QY 129 ELYQELTRIKAAVGLPEKSKLOEYQELTRIKAAVGLPEKSKLOEYQELTDLKIAPE 188

Db 117 AVHSEML-----LRVQQLVDDLKLTGQVATLNNNEE-----ASTE 153

QY 189 KICRHCKDKMTFFQGNFYFMSNSQRNMHDSVTAQCEVRKQLVVYIKTAQESNQLQOTERS 248

Db 154 GTC--GCVNNVNEHDSQYWFHSQSMALAEKRYGCLNHLVINSREKQNFVQKYL--G 209

QY 249 NRESWMLSDLNQGTQWQVWVDSPLSPFQRYVNSGEPNN-----STNFFPAEFSGSG- 301

Db 210 SAYTMMRLSD--PRGAKWVETIIVVATIDQGN-WKPEGLQWQWQGNH QXQDPE
QY 302 WNDNKQVDNWMICK 316
Db 267 WNDVCOBPHWCE 281

RESULT 14

US-09-535-521-20
Sequence 20: Application US/0955521

Patent No. 6410714

GENERAL INFORMATION:

APPLICANT: Weber, Eric R.

APPLICANT: McCall, Catherine A.

TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IGE RECEPTOR

FILE REFERENCE: AL-5

CURRENT APPLICATION NUMBER: 09/095,521

CURRENT FILING DATE: 2000-03-24

EARLIER APPLICATION NUMBER: 66/7125,913

EARLIER FILING DATE: 1999-03-24

NUMBER OF SEQ ID NOS: 26

SOFTWARE: patentin Ver. 2.1

SEQ ID NO 20

LENGTH: 208

TYPE: PRT

ORGANISM: Canis familiaris

US-09-535-521-20

Query Match 17.5% Score 296.5; DB 4: Length 208.
Best Local Similarity 31.5%; Pred. No. 8.5e-20;
Matches 68; Conservative 42; Mismatches 81; Indels

QY 123 EKSKLOEYQELTRIKAAVGLPEKSKLOEYQELTRIKAAVGLPEKSKLTQ 128

Db 1 QKSGAAVVSQIMKREIQAQKRM--KAQSELSSNLVATPLNNKLSQSLNI

QY 181 TDLTAFFERL-----CRHCKDKMTFFQGNFYFMSNSQRNMHDSVTA 248

Db 59 ERDQEVYKIMELHVSNGSEQNTPEKMLNPKKQVYRGEPRKMIQAFPA

QY 230 VVITAEQNFQQTQSRNFSNMRI SQNGESTWAWVGSPLSPFYK 292

Db 119 ASHSOEQDFLARYANKG--TWIGKQDREGEFIMMENPLNS--NK

QY 290 G-NEKAEFSGSG-WNDNRG--IVINWYTKKPAAG 321

Db 174 GQGFIVWMCSSQWNCADFSSSLQGWTEPLATG 208

RESULT 15
5514582-11
Patent No. 5514582
APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.
TITLE OF INVENTION: RECOMBINANT JMA ENVYING HYBRID
IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 43
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 09/097,187,670
FILING DATE: 21-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 986,941
FILING DATE: 08-DEC-1992
APPLICATION NUMBER: 808,122
FILING DATE: 16-DEC-1991
APPLICATION NUMBER: 440,624
FILING DATE: 22-NOV-1989
APPLICATION NUMBER: 315,015
FILING DATE: 23-FEB-1989
SEQ ID NO. 11.
LENGTH: 111
5514582-11

GenCore version 5.1.1
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2b model

Run on: December 7, 2002 11:20:21 Search time: 66 seconds

(without alignments)
1779 820 Million coll updates/sec

Title: US-09-831-458a-12

Perfect score: 1696

Sequence: 1 MDSKEPPVQGLIICGTH

Scoring table:

AltoSim2

Search: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 883724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-frame.p2b.model -DEV=xib
-O/cg2_1/705770.spool/US09831458/rnall_05122002_10x10_1510x7gpf-perry.fasta_1.519
-DB-Issued-Patents_NA -QFMT-laslap -SUFFIX-std.rml -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-bioSUM62 -TRANS-human40.cdi
-LIST=45 -DOCCALIGN=200 -THR SCORE-pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE-LOCAL -OUTFMT-p2b -NORM-ext -HEAVISIZE=500 -MINLEN=0 -MAXLEN=200000000
-USFR-US09831458.acncl_1117-strunc_05122002_103216_15169 NCTU-6 -ICPU-3
-NO_XLXAY -NO_MAP -LARGEQUERY -NPG_SCORES=0 -WAIT=1 -OWNCL=0 -REV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAP=10 -XGAPEXT=0.5 -FGAP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

1: Issued-Patents_NA:*
2: /cg2_6/prodata/1/ina/5A.COMB.seq:*
3: /cg2_6/prodata/1/ina/5B.COMB.seq:*
4: /cg2_6/prodata/1/ina/6A.COMB.seq:*
5: /cg2_6/prodata/1/ina/6B.COMB.seq:*
6: /cg2_6/prodata/1/ina/6C.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1412.5	83.3	1212	4	US-09-591-435-9
2	1412.5	83.3	1312	4	US-09-517-605-1
3	1407	83.0	1407	4	US-09-517-605-16
4	1398.5	82.5	1212	4	US-09-591-435-10
5	1393.5	82.2	1312	4	US-09-591-435-11
6	3350	20.6	1005	1	US-08-365-1038-1
7	338	20.0	924	1	US-08-365-1038-5
8	335	19.8	885	1	US-08-365-1038-3
9	329.5	19.4	876	4	US-09-535-521-4
10	329.5	19.4	876	4	US-09-535-521-6
11	329.5	19.4	2851	4	US-09-535-521-1
12	329.5	19.4	2851	4	US-09-535-521-3

13	322.5	19.0	1047	1	US-08-465-1038-7
14	319.5	18.8	1075	1	US-08-465-1038-9
15	319.5	18.8	1458	4	US-09-111-470-3
16	305	18.0	1370	4	US-09-111-470-4
17	296.5	17.5	624	4	US-09-535-521-19
18	296.5	17.5	624	4	US-09-535-521-21
19	290.5	16.7	561	4	US-09-535-521-16
20	284.5	16.7	561	4	US-09-535-521-18
21	263	15.5	417	4	US-09-535-521-10
22	263	15.5	417	4	US-09-535-521-12
23	263	15.5	423	1	US-09-535-521-13
24	263	15.5	423	4	US-09-535-521-15
25	256	15.1	384	4	US-09-535-521-7
26	256	15.1	384	4	US-09-535-521-9
27	246	14.5	1318	4	US-08-809-494A-5
28	246	14.5	1318	4	US-08-809-494A-7
29	243.5	14.4	369	4	US-09-535-521-24
30	243.5	14.4	369	4	US-09-535-521-26
31	232	13.7	1104	4	US-09-111-470-1
32	230	13.6	501	1	US-09-111-470-1
33	230	13.6	1227	4	US-08-772-440-20
34	228	13.4	394	4	US-08-772-440-39
35	226.5	13.4	1418	4	US-09-111-470-7
36	225.5	13.3	1887	4	US-08-809-494A-1
37	225.5	13.3	1887	4	US-09-452-402-1
38	225.5	13.3	1906	4	US-08-809-494A-3
39	225.5	13.3	1906	4	US-09-452-402-3
40	216.5	12.8	402	4	US-08-543-246B-10
41	216.5	12.8	648	4	US-08-543-246B-14
42	216.5	12.8	1755	4	US-08-543-246B-8
43	202.5	11.9	4771	4	US-08-840-062-1
44	200.5	11.8	4588	4	US-08-840-062-1
45	196.5	11.6	871	1	US-08-650-578-1

ALIGNMENTS

RESULT 1
US-09-591-435-9
Sequence 3, Application US/09591435
Patent No. 6280953
GENERAL INFORMATION:
APPLICANT: MESSEIER, WALTER
TITLE OF INVENTION: METHODS TO IDENTIFY POLYNUCLEOTIDE AN
TITLE OF INVENTION: SEQUENCES WHICH MAY BE ASSOCIATED W
FILE REFERENCE: GENO. 200.2
CURRENT APPLICATION NUMBER: 05/07591, 435
PRIOR APPLICATION NUMBER: 2000-06-09
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 09/541, 435
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 09/240, 915
PRIOR FILING DATE: 1999-01-29
PRIOR APPLICATION NUMBER: 60/074, 253
PRIOR FILING DATE: 1998-01-30
PRIOR APPLICATION NUMBER: 60/094, 987
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 1212
TYPE: DNA
ORGANISM: Homo sapiens
US-09-591-435-9
Alignment scores:
Pred. No.: 7.3e-145
Score: 1412.50
Percent Similarity: 77.04%
Best Local Similarity: 71.94%
Query Match: 83.28%
DB: 4
Gaps: 4

QY	121	LeupProgluylusSerLysIle	-----	127
Db	450	CTTCACAGCAATCTAAGCTCCAGAGATCTACTAAAGTCACTCGATGAAGGATGCA	-----	509
QY	127	-----	-----	127
Db	510	GTGATGACTTCCAGAGAAATTTAAGATGACAGAGATATTACAGAGATCACTGGGCTG	-----	569
QY	128	-----	-----	128
Db	570	AAGGTGTACTGGGTGGGTTTCCAGAGAAATTTAACTACAGACATCTATACAGATCTG	-----	629
QY	135	ThraFdeulysAlaIaValGlyIleuProGlyLysSerLysLeuGlnIleuThr	-----	154
Db	630	ACCCGGCTGCAAGGCTGCAAGTGGTTCAGACTTCAGAGAAATCTAAGACACGACAGATTAC	-----	689
QY	155	GlnGluLeuThraFdeulysAlaIaValGlyIleuProAspGlnIleuSerGln	-----	174
Db	690	CAGAGAGTGTACCTGGAGGTGTGAGATGGGTGTGACTTTTCAGAGAAATCTAAGCAACAG	-----	749
QY	175	GlnIleuThrGlnIleuLeuThrAspLeuLysThrAlaPheGluIleuGlyAsnArgIleu	-----	194
Db	750	GAGATCTACAGAGAGCTGACCCCACTGAAAGGCTGCAAGTGGAGAGGCTGTGGCACCCCTGT	-----	809
QY	195	ProLysAspTrpThrPhePheGlnGlyAsnGlySerPheMetSerAsnSerGlnAsn	-----	214
Db	810	CCCTGGCAATATGACATCTCTCCCAAGGAACCTGTACTCTATCTATCACTCCACGCGAGAC	-----	869
QY	215	TrpHisAspSerValThrAlaGlyGlnIleuValArgAlaGlnIleuValIleuLysThr	-----	234
Db	870	TGGACACACTCTCCATCCCGGCTGTAAAGAAAGTGGGGGCTGACTGTGTGTAATCAAAAGT	-----	929
QY	235	AlaGlnIleuGlnAsnPheLeuGlnIleuGlnIleuSerAspSerAsnArgPheSerTrpMet	-----	254
Db	930	GCTGAGAGAGCAAACTCTCCACAGCTGCAGACTCTCCAGAACATCAACCTTAACCTCAGT	-----	989
QY	255	GlyLeuSerAspLeuAsnGlnIleuGlyThrTrpGlnIleuValAspLysSerIleuSer	-----	274
Db	990	GGACTTTCAGCTTAATTCAGGAAGGCTACCTTCAATTCAGATCTGCAATCACTGCTGTTC	-----	1049
QY	275	ProSerPheGlnAspTrpTrpAsnSerGlyIleuProAsnAsnSerGlyAsnGlnuPds	-----	294
Db	1050	CCGACCTTCAGAGAGATTTGACACATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	-----	1109
QY	295	AlaGlnPheSerGlySerGlyTrpAsnAspAsnArgGlyAspValAspAsnIleuTrpIle	-----	314
Db	1110	GGCGAAATTTAGTGGCAATGTCTTAACATGACCAAAATTTAACTTGGCAAAATCTGGATC	-----	1169
QY	315	CysLysLysProAlaIaIleuGlyPheAspAspGln	-----	325
Db	1170	TGCAAAAAGTCCGACAGCTTCTTCTTCCAGGAGTAA	-----	1205
RESULT 3				
US-09-517-605-16				
: Sequence 16, Application US/09517605				
: Patent No. 6391567				
: GENERAL INFORMATION:				
: APPLICANT: Litchman, Dan R.				
: APPLICANT: Kwon, Douglas S.				
: APPLICANT: van Kooyk, Yvette				
: APPLICANT: Geltenbeck, Theo				
: TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO				
: TITLE OF INVENTION: CELLS				
: FILE REFERENCE: 1049-1-017				
: CURRENT APPLICATION NUMBER: US/09/517,605				
: CURRENT FILING DATE: 2000-03-02				
: NUMBER OF SEQ ID NOS: 17				
: SOFTWARE: PatentIn Ver. 2.0				
: SEQ ID NO 16				
: LENGTH: 1643				
: TYPE: DNA				
: ORGANISM: Homo sapiens				
: FEATURE:				

:	NAME/KEY:	CDS
:	LOCATION:	(25)
:	US-09-517-605-16	(915)

Alignment Scores:		
Pred.: No.:	4,526-144	Length:
Score:	1407.08	Matches:
Percent Similarity:	76.88%	Conservative:
Best Local Similarity:	76.88%	Mismatches:
Query Match:	82.96%	Indels:
RE:	4	Gaps:
		2

US-09-531_458A-12 (1,325) x US-09-517_605-16 (1,1643)

OY	1	MetSerAspSerLysGlnArgValHisIleLeuGlyLeuLeuGlyCys	
Dd	25	ALGNAGTAAATTAAGAATAAAGAGAGAAATGAATTAAGATTCTGGTCCTC	
OY	21	GlyAlaLeuValLeuGlnLeuLeuSerPheMetLeuLeuAlaGlyValLeu	
Dd	85	GACCCCGCTGTCTGCACACTGCTCTCTCTCAAGCTCTTGCTGGCTGCTG	
OY	41	LeuValGlnValSerLysValProSerSerLeuSerGlnIleGlnSerGln	
Dd	145	CTGTCCAAAGTGTCCAAAGGTGCCAGATCCCATTAAGTACAGAACAAATCCC	
OY	61	IleTyrGlnAsnLeuThrGlnLeuLysAlaAlaValGlyLeuSerGln	
Dd	205	ATCTACCAAGACCTGCACCAGCTCAATACCTGCACCTGGTCAAGTCTCACAA	
OY	81	LeuGlnGlnIleTyrGlnGlnIleThrGlnLeuLysAlaAlaValGly	
Dd	205	CTCTAAACAATCAATAGAAATTAATTAATTAATTAATTAATTAATTAATTA	
OY	101	LysSerLysLeuGlnIleGlnTyrGlnGlnLeuThrArgLeuLysAlaAla	
Dd	325	AAGTTAAATTATAGTAAGATTAATTAAGAAATTAATTAAGAAATTAATTA	
OY	121	LeuProGlnLysSerLysLeuGlnIleIleTyrGlnGlnLeuThrArg	
Dd	385	TTCGCAAGAAATACCAAGGTGTACAGAAATCAACCAAGCTGCACCGCTCT	
OY	141	ValGlyIleIleLeuProGlnLysSerLysLeuGlnIleIleTyrGlnLeu	
Dd	445	GAAATTAAGTTATAGTAAGAAATTAATTAATTAATTAATTAATTAATTA	
OY	159	-----	
Dd	505	AAGGCTACATGGGTGAAGTACCAATAAATACCAAGCTGCAGCAATCTTA	
OY	159	-----	
Dd	565	AGCGAGTTGAAGCTGCAGTAAATTAATTAATTAATTAATTAATTAATTA	
OY	160	-----LeuLysAlaAlaValGlyIleLeuProAspIleGln	
Dd	625	CATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	
OY	175	GlnIleTyrGlnGlnLeuThrAspLeuLysThrAlaPheLeuThrValGly	
Dd	685	CAAACTCATCAAGCAATGCATTTAATAGCTGCATTTCAACGGCTGCTG	
OY	195	ProLysAspPrpThrPheLeuGlnIleLysAsnGlySerPheMetSerAsn	
Dd	745	CCCAAGACCTGACATCTCTCAAAAATTAATTAATTAATTAATTAATTA	
OY	215	TrpHisAspSerValThrAlaGlyGlnIleValArgAlaGlnIleLeuVal	
Dd	805	TGGAGTAGCTGGTCAACCTCTGCTAGGAATTAAGAGGCTCAGCTGCTG	
OY	235	AlaGlnGlnGlnAsnPheLeuGlnIleLeuGlnIleThrSerArgSerAsnArg	
Dd	865	GCTGAGGAG-----	

Alignment Scores:					
Pred No	2	46-143	Length:	1212	
Score:	1398.50		Matches:	277	
Percent Similarity:	76.53%		Conservative:	24	
Best Local Similarity:	70.66%		Mismatches:	21	
Query Match	82.46%		Indels:	71	
DH:	4		Gaps:	4	

US-09-831-458A-12 (1-325) x US-09-591-435-10 (1-1212)		
QY	1 MetSerASPScRlySLySuhTAR+ValGICInLeuCYLeuCG-----	15
Ib6	1 ATAGAGTACTCAAGSAACAAACAAGCTACAGCANTGGCTGTGTGAGGASAAACATG	60
QY	16 -----GLYCylscLgThIs 20	
Ib6	61 AGANSCCTTCGATTCCGAACACACGAGCGCTAACAGAGCTACAGCGCTTCCTGCAT	120
QY	21 AAlafenValIcglInleucAsrPhAcLdeuAlaLyVAlLeuVALAlAlle 40	
Ib6	121 GGRGCTCTATGTATTAATTTCTCTTACAGCTTTTGATGAGG-----	160
QY	41 LeuValClIvalScRlySLyValProSeSerLeuSerGlnoIdnSerGIuIdnAsPala 60	
Ib6	169 GTTAGCAAGATGTCAAAGCTTGTTCAGCTDVAIAAGTCAAGSAACATCAAGTAAACAGTGC 220	
QY	61 lletyrGIhaSnLeuThrIlLeuYSAlAIValcLYlInleuSerGIuysSerlys 80	

2b	ATGAAAGCAAGTTCAGAGAGGCTTAAAGGCTTAAATGGATGGCTCTACAGAAATTCAGAG	288
QY	LeuGlnGluLeuLeuArgGlnGlnLeuThrGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu	100
Db	CTGATGAGAGATGATGAG	348
QY	LeuSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu	120
Db	AAATGCTAAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	408
QY	LeuProGlnLeuSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu	140
Db	CTGATGAG	468
QY	ValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	160
Db	CTGATGAG	528
QY	LeuValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu	180
Db	AAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	588
QY	ThrAsp-----	182
Db	AG	648
QY	182-----	182
Db	ATGAG	708
QY	183-----LeuLeuThrAlaPheGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	194
Db	AG	768
QY	ProGlnSerPhePhePheGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	214
Db	AG	828
QY	TrpHisAspSerValThrAlaCysGlnGlnValAlaGlnGlnLeuValValIleLeuThr	234
Db	AG	888
QY	AlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	254
Db	AG	948
QY	GlyLeuSerAspLeuAsnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	274
Db	AG	1008
QY	ProSerThrLeuAlaTrpTrpPheSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	294
Db	AG	1068
QY	AlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	314
Db	AG	1128
QY	GlyLeuSerTrpAlaValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	325
Db	AG	1164
RESULT 5		
US-89-591-4-45-11		
Sequence 11, Application US-09591435		
Patent No. 5,280,954		
GENERAL INFORMATION		
APPLICANT: MESSIER, WALTER		
APPLICANT: SIKELA, JAMES M		
TITLE OF INVENTION: METHODS TO IDENTIFY POLYNUCLEOTIDE AND PEPTIDE		
TITLE OF INVENTION: SEQUENCES WHICH MAY BE ASSOCIATED WITH PHYSIOLOGICAL		
AND MEDICAL CONDITIONS		

[illegible]


```

QY 46 LysValProSerSetLeuSerGlnGlnGlnSerGlnGlnAspAlaLeuTyr----- 62
Db 3 AAGCATCGAAGCAAGACTGGATATGAGAAATGAAATACACATGAGAGGATGAG 62
QY 63 -----GlnAsnLeuThrGlnLeu-----Tyr 69
Db 63 CTGGCGCTCTCTTTTCTGTTGTCATCTGGAAACGGAAGACATCTAAACACCTGGAGAC 122
QY 70 AlaAlaValGlyLeuSerGlnTyrSerLeuGlnLeuTyrGlnGlnLeuThr 89
Db 123 ACTGCAATTCAGATCTCTCTCATCTTACCAATGACCTACAAAATTCAGACT----- 176
QY 90 GlnLeuLysAlaAlaValGlyLeuLeuProGlnTyrSerTyrLeuGlnLeuTyrGln 109
Db 177 -----AAATCAATTCGCGAAGATCTGAAATCTGAAATCTGTTTTCAGATCTGACAA 218
QY 110 GlnLeuThrArgLeuLysAlaAlaValGlyLeuLeuProGlnTyrSerTyrLeuGln 129
Db 219 AACTTCAGAGACTCGAAGCTGAACATCAACAAATG-----AAAGCTAGGACTCTGG 272
QY 130 LLeuTyrGlnLeuLeuThrArgLeuLysAlaAlaValGlnGlnLeuLeuTyrSerTyr 149
Db 273 CTCTCCCAAGCAAGCTGACCGGACTCCAGGAGATCTAAGCAAC-----GCCCATCCGAG 326
QY 150 LeuGlnGlnLeuTyrGlnGlnLeuThrArgLeuLysAlaAlaValGlyLeuLeuProGln 169
Db 427 AACTCAAAACTCTCCCGAAGCTGAAATCTGAAACATCTGAAACATCTGAAATCTGAA 386
QY 170 -----GlnSerTyrGlnGlnGlnGln 176
Db 487 CTGGGCTTGAATGAGAGCGGACAGCCCTCCGATTCTCTACAGAACTCCAGCAAGAGG 446
QY 177 TyrGlnLeuLeuThrArgLeuLysTyrAlaPheGlnArgLeuSerThrLeuTyrGln 196
Db 447 GCAAGCTGTGATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 506
QY 197 AspThrPhePheGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 216
Db 507 AACTGGCTCCATTCGCAACAAATGCTTACTATTCTTCTGCAAGGCTCCGCAAGCTGATG 566
QY 217 AspSerValThrAlaGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 236
Db 567 CAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 626
QY 237 GlnGlnAsnPheLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 256
Db 627 GAACAGGACTCTCTGATGCAACATCAACATCAACATCAACATCAACATCAACATCAAC 680
QY 257 SerAspLeuAsnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 276
Db 681 CAGCATCTCAATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 740
QY 277 PheGlnArgTyrTrpAsnSerGlnGlnProAsnAsnSerGlnGlnGlnGlnGlnGln 295
Db 741 -----AACTGGAATTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 791
QY 296 GlnPheSerGlySerGlyTyrTrpAsnAspAsnArgCys-----AspValAspAsnTyrTrp 313
Db 792 ATGATCGCGGATCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 851
QY 314 LLeuGlyLysLysProAlaAlaCys 321
Db 852 GTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 875

```

```

TITLE OF INVENTION: DNA Sequences for Soluble Forms of
TITLE OF INVENTION: and Methods of Use for Same
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESS: Zartley, McKee, Thumby, Voorhees & Sease
STREET: 901 Grand Ave. Suite 4200
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPILED: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-09/465,103B
FILING DATE: 28-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Nebel, Heidi S.
REGISTRATION NUMBER: 37,719
REFERENCE/PRIOR ART NUMBER: 5111, 25-24
TELEPHONE: (515) 288-3667
TELEFAX: (515) 288-1338
INFORMATION FOR SEQ. ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 885 base pairs
TYPE: nucleic acid
STRANDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 24..884
US-08-365-103B-3
US-09-831-458A-12 (1-325) x US-08-365-103B-3 (1-885)
Alignment Scores:
Pred. No.: 1,616 27 Length: 885
Score: 335.00 Matches: 94
Percent Similarity: 48.21% Conservative: 57
Best Local Similarity: 29.64% Mismatches: 99
Query Match: 19.75% Indels: 60
DB: 1 Gaps: 14
QY 46 LysValProSerSetLeuSerGlnGlnGlnSerGlnGln----- 62
Db 3 AAGCATCGAAGCAAGACTGGATATGAGAAATGAAATACACATGAGAGGATGAG 62
QY 59 -----AspAlaLeuTyrGlnAsnLeuThrGlnLeu----- 69
Db 63 AATCTAAACAGCTGGAGAG/CTGCAATTCAGAAATGATCTCTCATCTATA----- 122
QY 69 ---LysAlaAlaValGlyLeuSerGlnTyrSerLeuGlnLeuTyrGlnGlnLeuThr 89
Db 123 CAAAATTCAGACTAAATCAATTCGCGAAGCTGAAATCTGAAATCTGTTTTCAGATCTGACAA 218
QY 88 LeuThrGlnLeuLysAlaAlaValGlyLeuLeuProGlnTyrSerTyrLeuGln 129
Db 183 TTTCAGAGACTCGAAGCTGAACATCAACAAATG-----AAAGCTAGGACTCTGG 272
QY 108 TyrGlnGlnLeuLeuThrArgLeuLysAlaAlaValGlyLeuLeuProGln 149
Db 237 TCCAGAGACTCGAAGCTGAACATCAACAAATG-----AAAGCTAGGACTCTGG 272
QY 128 GlnGlnLeuTyrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 196
Db 291 TCAAAACTCTCCCGAAGCTGAAATCTGAAACATCTGAAACATCTGAAATCTGAA 386

```


[illegible]

•
•
•
•

GenCore version 5.1.4
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2002, 11:18:56, Search time 61 Seconds
(without alignments)
2077.510 Million cell updates/sec

Title: US-09-831-458A-12

Perfect score: 1696 1 MCSKSEPRVQGLGILGTHH PRTVWVWVTKRPAPCPPEPP 325

Scoring table:

BLOSUM62	
Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 350425 seqs, 194966369 residues

Total number of hits satisfying chosen parameters: 700950

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_model -DEV=xlh
-O=/cgn2_1/USPRN/spool/US09831458/rnat_05122002_109217_15180/aff_query_fast4_1_519
-DB=Published.Applications_NA -QMT=fastap -SUFFIX=std.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THP_SCOPE=opt -THP_MAX=100
-THP_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -HAPPSTZ=500 -MINLEN=0
-MATEL=2000000000 -USPR=US09831458 -CGN=1_1_21 -rnat_05122002_109217_15180
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MMAP -TAREPQUERY -NPG_SCOPE=0 -WAIT -TONTLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/prodata/2/pubpna/PTCT_NFM_PUB.seq.*
3: /cgn2_6/prodata/2/pubpna/US06_NFM_PUB.seq.*
4: /cgn2_6/prodata/2/pubpna/US06_NFM_PUB.seq.*
5: /cgn2_6/prodata/2/pubpna/US07_NFM_PUB.seq.*
6: /cgn2_6/prodata/2/pubpna/US08_NFM_PUB.seq.*
7: /cgn2_6/prodata/2/pubpna/US08_NFM_PUB.seq.*
8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/prodata/2/pubpna/US09_NFM_PUB.seq.*
10: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/prodata/2/pubpna/US10_NFM_PUB.seq.*
12: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/prodata/2/pubpna/US10_NFM_PUB.seq.*
14: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	324	19.4	1617	10	US-09-728-952-45
2	194	18.8	1458	9	US-09-862-802-3
3	305	18.0	1370	9	US-09-862-802-9
4	296	17.5	1355	9	US-09-878-295A-230

Result No.	Score	Query Match	Length	DB ID	Description
5	296.5	17.5	1355	9	US-09-878-697-230
6	296.5	17.5	1355	9	US-09-878-192A-230
7	289	17.0	592	10	US-09-864-763-7450
8	284	16.7	152	10	US-09-864-761-24072
9	275.5	16.2	1277	10	US-09-889-107-2230
10	274.5	16.2	2099	10	US-09-764-820-14
11	273.5	16.1	2099	10	US-09-905-291A-376
12	273.5	16.1	997	10	US-09-909-330-376
13	273.5	16.1	997	10	US-09-909-088A-376
14	273.5	16.1	997	12	US-10-052-586-23
15	267.5	15.8	1300	10	US-09-880-107-4731
16	267	15.7	1309	10	US-09-880-107-4235
17	263.5	15.5	549	12	US-10-090-465-3
18	263.5	15.5	642	12	US-10-090-465-3
19	263	15.5	3750	10	US-09-917-800A-474
20	253.5	14.9	1036	10	US-09-843-381-119
21	249.5	14.7	758	10	US-09-843-381-1335
22	246.5	14.5	1290	10	US-09-917-800A-1642
23	234.5	13.8	3763	10	US-09-870-759-141
24	232	13.7	1096	10	US-09-764-870-199
25	232	13.7	1096	10	US-09-764-870-199
26	232	13.7	1104	10	US-09-962-802-1
27	230	13.6	2940	10	US-09-745-763-198
28	226.5	13.4	1418	10	US-09-862-802-7
29	225	13.3	1209	12	US-10-052-586-593
30	215	12.7	693	10	US-09-843-381-994
31	211.5	12.5	443	10	US-09-960-352-7063
32	210.5	12.4	5185	10	US-09-870-759-94
33	206	12.1	759	10	US-09-944-807-5
34	202.5	11.9	562	10	US-09-744-870-124
35	202.5	11.9	1066	10	US-09-744-870-15
36	201	11.9	528	10	US-09-843-381-995
37	201	11.9	743	10	US-09-764-870-16
38	195	11.5	1238	10	US-09-992-598-356
39	195	11.5	1238	10	US-09-992-598-356
40	195	11.5	1238	10	US-09-989-723-356
41	195	11.5	1238	10	US-09-989-723-356
42	195	11.5	1238	10	US-09-989-723-356
43	195	11.5	1238	10	US-09-989-723-356
44	195	11.5	1238	10	US-09-989-723-356
45	195	11.5	1238	10	US-09-989-723-356

ALIGNMENTS

RESULT 1
US-09-728-952-45
Sequence 45, Application US/09728952
Parent No. US2002011102A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Wang, Jian-Rui
APPLICANT: Wang, Dunrui
APPLICANT: Yamazaki, Yuki
APPLICANT: Ujwal, Manusha L.
APPLICANT: Drmanac, Radovic L.
TITLE OF INVENTION: No. US2002011102A1 Nucleic Acids
FILE REFERENCE: Polypeptides
CURRENT APPLICATION NUMBER: US20020728952
CURRENT FILING DATE: 2000-11-10
NUMBER OF SEQ ID NOS: 101
SOFTWARE: pL_genes Version 2.0
SEQ ID NO 45
LENGTH: 1617
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS

APPLICANT: Gerber, Hanspeter
APPLICANT: Gottisen, Mary F.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavan, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Thomas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William L.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630P1C11
CURRENT APPLICATION NUMBER: US/09/978,295A
PRIOR FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/9418585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/004249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/005411
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/06364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/07632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078946
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078940
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078949
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079456
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079554
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079746
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31

PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080344
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082559
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082894
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742

[illegible]

```

01  Oy      303  DASHASUAGQYASAPVJASP    ASUTITRIPLETYSLSYSTFOT
02  Db       888  CGACGACCGCTGTGTGCAGS QVAMAAZACGACGTCGA
03                                     IIII   IIIII   :::   :::   IIIIII:IIIIII
04                                     :::   :::   :::   :::   :::   :::
05 RESULT 5
06 US-09-978--697-230
07 Sequence 230. Application US/09978697
08 Patent No. 052852C15928A1
09 GENERAL INFORMATION:
10 APPLICANT: Ashkenazi, Avi
11 APPLICANT: Baker Kevin P.
12 APPLICANT: Betsteijn, David
13 APPLICANT: Desnoyers, Luc
14 APPLICANT: Eaton, Dan
15 APPLICANT: Ferrara, Napoleon
16 APPLICANT: Filvarotti, Elton
17 APPLICANT: Fong, Sherman
18 APPLICANT: Gao, Wei-Qiang
19 APPLICANT: Geibler, Hanspeter
20 APPLICANT: Gerilstein, Mary E.
21 APPLICANT: Goddard, Audrey
22 APPLICANT: Godowski, Paul J.
23 APPLICANT: Grimaldi, J. Christopher
24 APPLICANT: Gurney, Austin L.
25 APPLICANT: Hillan, Kenneth J.
26 APPLICANT: Kiljavin, Ivar J.
27 APPLICANT: Kuo, Sophia S.
28 APPLICANT: Nefter, Mary A.
29 APPLICANT: Pan, James:

```

APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William J.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630PIC27
CURRENT APPLICATION NUMBER: US/09/078,607
PRIOR FILING DATE: 2001-10-17
PRIOR APPLICATION NUMBER: 09/91,8585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077652
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078946
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078940
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078949
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-04-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079693
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333

PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081917
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081919
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081938
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083346
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083445
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07

[illegible]


```

: CURRENT FILING DATE: 2001-06-14
: PRIOR APPLICATION NUMBER: US 60/211,379
: PRIOR FILING DATE: 2000-06-14
: PRIOR APPLICATION NUMBER: US 60/237,054
: PRIOR FILING DATE: 2000-10-02
: NUMBER OF SEQ ID NOS: 3050
: SOFTWARE: Patent Ver. 2.1
: SEQ ID NO: 2230
: LENGTH: 1277
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: Genbank Accession No. U520020142082A1 M10053
US-09-831-458a-12

```

```

Alignment Scores:
Pred. No.: 5,31e-21 Length: 1277
Score: 275.50 Matches: 91
Percent Similarity: 39.60% Conservative: 46
Best Local Similarity: 26.30% Mismatches: 114
Query Match: 16,248 Indels: 95
DB: 10 Gaps: 14

```

US-09-831-458a-12 (1-325) x US-09-880-107-2230 (1-1277)

```

QY 6 GlnProArgValGlnGlnLeuGlyLeuLeuGlyLeuLeuGlyLeuValLeu 25
DB 257 CAGGCTCTCTCTGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 304
QY 26 GlnLeuLeuSerPheMetLeuLeuAlaGlyValLeuValAlaLeuValSer 45
DB 305 TCCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 343
QY 46 LysValIleProSerLeuSerLeuGlnGlnGlnSerGlnGlnAlaIleTyrGlnLeu 65
DB 344 GTGATCGATCCCAAACTCCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 403
QY 66 ThrGlnLeuLysAlaValAlaGlyLeuSerGlyLeuSerGlyLeuGlnIleTyr 85
DB 404 AGCAACTTCACAGCGACGACGACGACGACGACGACGACGACGACGACGACGAC 427
QY 86 GlnGlnLeuThrGlnLeuLysAlaValAlaGlyLeuProGlyLeuSerLeuGln 105
DB 428 GCCCAGGTCAAGGCTTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 487
QY 106 GluIleTyrGlnGlnLeuThrArgLeuLysAlaValAlaGlyLeuProGlyLeuSer 125
DB 488 GAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 493
QY 126 LysLeuGlnGlnIleTyrGlnGlnLeuThrArgLeuLysAlaValAlaGlyLeuPro 145
DB 494 CAGCTGAGAAATAGTCAAAATAGTCAAAATAGTCAAAATAGTCAAAATAGTCA 526
QY 146 LysLeuSerGlyLeuGlnGlnIleTyrGlnGlnLeuThrArgLeuLysAlaValAlaGly 165
DB 527 GATCACTCCACCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 568
QY 166 GluLeuProAspGlnSerLeuGlnGlnGlnIleTyrGlnGlnLeuThrAspLeuSer 185
DB 569 GACTGGGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 613
QY 186 AlaPheGlnAlaGlnLeuGlyLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 205
DB 614 GAGCTCAAAAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 667
QY 206 TyrPheMetSerAspSerGlnArgAspTrpHisAspSerAlaThrAlaGlnGlnVal 225
DB 668 TATGGTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 727
QY 226 ArgAlaGlnLeuValAlaIleTyrThrAlaGlnGlnGlnGlnGlnGlnGlnGlnGln 245
DB 728 GAGCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 787

```

```

QY 246 SerArgSerAsnArgPheSerTrpMetIleLeuSerAspLeuAsnIleLeu 265
DB 788 GCGCCTGTGAGC-----ACCTGATGAGGCTGTGAGC-----GAAAG 265
QY 266 GluTrpValAspCysSerProLeuSerTrpSerThrAlaArgTyrTrpAsp 285
DB 846 AAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 285
QY 286 TrpAsnAsn-----SerGlyAsnIleAspCysAlaGln 305
DB 893 CCGGACGACTGTACGGCAAGGATGAGATGAGATGAGATGAGATGAGATGAGAT 305
QY 300 SerCly TrpAsnAspAspAlaTyrAlaValAspAsnTyrTrpIleGly 325
DB 953 CACGCGCGCTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 325
QY 317 ----- 345
DB 1013 CTGCATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 345
QY 319 AlaAlaCysPheArgAsp 364
DB 1073 GCGGCAAGGCTCGGCAAT 1098
RESULT 10
US-09-764-870-14
: Sequence 14, Application US/09764870
: Patent No.: US,200,004,2486A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibody
: FILE REFERENCE: PT214
: CURRENT APPLICATION NUMBER: US/09764,870
: PRIOR APPLICATION DATA REMOVED - CONSULT PALM OR FILE WR
: NUMBER OF SEQ ID NOS: 646
: SOFTWARE: Patent In Ver. 2.0
: SEQ ID NO 14
: LENGTH: 2099
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-764-870-14
Alignment Scores:
Pred. No.: 1.39e-20 Length: 2099
Score: 274.50 Matches: 91
Percent Similarity: 44.30% Conservative: 51
Best Local Similarity: 27.18% Mismatches: 124
Query Match: 16,198 Indels: 44
DB: 10 Gaps: 12
US-09-831-458a-12 (1-325) x US-09-764-870-14 (1-2099)
QY 36 ValLeuValAlaIleLeuValGlnValSerGlyValProSerSerLeuSer 325
DB 174 CTGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 325
QY 56 SerGlnGlnAspAlaIleTyrGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 345
DB 234 GAT-----GTAAAGATGATGATGATGATGATGATGATGATGATGATGAT 345
QY 76 SerClyLysSerLeuGlnGlnIleTyrGlnGlnGlnGlnGlnGlnGlnGlnGln 365
DB 282 ACGACCGCTGAT-----TCTGAATTAATAATAATAATAATAATAATAATA 365
QY 96 GlyIleGlnGln-----GlnLysSerLysLeuGlnGlnIleTyrGlnGln 385
DB 336 GTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 385
QY 115 LysAlaAlaValAlaGlyLeuProGlnIleLysSerLysLeuGln----- 405
DB 396 AAACGACGTGTG-----GAGAAAGCTGATGATGATGATGATGATGATG 405

```

[illegible]

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Tue Dec 10, 10:10:15 2002

us-09-831-458a-12.std.rnpb

Page 17

Db 851 AATGGCGATGGCAACAGTATTATGGCAAAATATCAAAAC---TGGGTGTGATCG 907

Gy 286 ProAsnAsn-----SerGlyAsnGlyAspCysAlaGluPheSerGly 299

Db 908 CGAGATATATGGCGACGGGACAGATGGGTGGGTGAAAGCTGAGAGACTGTGTGAAGTGCACCG 967

Gy 300 SerGly---TrpAsnAspAsnAlaGlyAspValAspAsnIYTrpIleGlyLysLysPro 318

Db 968 GATATGCTGTGAACATATATTTTACTGTAAGCTTGGGTGTGTGTGATGAAAG- 1026

Gy 419 AlaAlaCysPheArg 323

Db 1027 GGCGAATGCCACCGG 1041

Search completed: December 7, 2002, 11:29:28
Job time : 76 secs

...

Genforce version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: December 7, 2002, 09:48:14 : Search time 26 seconds

(without alignments)
518.454 Million cell updates/sec

Title: US-09-831-458A-12

Perfect score: 1696

Sequence: 1 MSDSKEPPVQGLGLDGLSH.....PQVDNMYLCKKPAACPFDE 325

Scoring table: HUSUM62

Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_40.*

Prod No is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query length	DB ID	Description
1	350	20.6	331	1	PCB2_MOUSE
2	326.5	19.3	207	1	LECH_CHICK
3	322.5	19.0	321	1	PCB2_HUMAN
4	309.5	18.2	550	1	KURC_RAT
5	304	18.2	283	1	LECH_MOUSE
6	307	18.1	304	1	MMGL_MOUSE
7	289	17.0	283	1	LECH_MOUSE
8	287	16.9	306	1	MMGL_RAT
9	284.5	16.8	548	1	KURC_MOUSE
10	271.5	16.0	290	1	PCB2_MOUSE
11	267	15.7	311	1	LECH_HUMAN
12	258	15.2	301	1	LECH_MOUSE
13	240	14.2	301	1	LECH_MOUSE
14	234.5	13.2	216	1	NECD_MOUSE
15	219.5	12.9	162	1	LECH_MOUSE
16	216.5	12.8	216	1	NECD_MOUSE
17	210.5	12.4	146	1	MANE_HUMAN
18	206	12.1	149	1	CLT2_HUMAN
19	202.5	11.9	223	1	NK13_RAT
20	202.5	11.9	2415	1	PCB2_HUMAN
21	201.5	11.9	2124	1	PCB2_MOUSE
22	199	11.7	883	1	PCB2_MOUSE
23	197.5	11.6	179	1	PCB2_MOUSE
24	197	11.6	2132	1	PCB2_MOUSE
25	196.5	11.6	179	1	PCB2_MOUSE
26	195.5	11.5	179	1	PCB2_MOUSE
27	195	11.5	883	1	PCB2_MOUSE
28	195	11.5	2109	1	PCB2_MOUSE
29	194.5	11.5	912	1	PCB2_MOUSE
30	194.5	11.5	2333	1	PCB2_MOUSE
31	193	11.4	2364	1	PCB2_MOUSE
32	191	11.3	199	1	PCB2_MOUSE
33	189.5	11.2	1268	1	PCB2_MOUSE

34	188	11.1	227	1	NK11_MOUSE
35	186.5	11.0	1257	1	PCB2_MOUSE
36	180.5	10.6	147	1	LECH_MOUSE
37	178	10.5	244	1	MANE_MOUSE
38	176.5	10.4	3562	1	LECH_MOUSE
39	175	10.3	612	1	LECH_MOUSE
40	174	10.3	175	1	ANP_MOUSE
41	171	10.1	223	1	NK12_MOUSE
42	170	10.0	165	1	LECH_MOUSE
43	169	10.0	125	1	ARAB_MOUSE
44	168	9.9	2738	1	PCB2_MOUSE
45	168	9.9	3358	1	PCB2_MOUSE

ALIGNMENTS

RESULT 1	ID	PCB2_MOUSE	STANDARD	PROT	AA
AC	130633	Q61556	Q61557		
DT	01-FEB-1991	(Rel. 17, Created)			
DT	01-FEB-1991	(Rel. 17, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Low affinity immunoglobulin epsilon FC receptor (lympho-				
DE	receptor) (FC-epsilon-R1) (Clz).				
GN	FCER2 OR FCER2A				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Eute-				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mu-				
OX	NCBI Taxid=10090				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=90017519; PubMed=2529542				
RA	Bettler B., Holstetter H., Kao M., Yokoyama W.M., Krich				
RT	Conrad D.H.				
RT	"Molecular structure and expression of the murine lympho-				
RT	affinity receptor for IgE (Fc epsilon R1)."				
PI	Proc Natl Acad Sci U S A 87:7565-7570(1989).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=90171598; PubMed=217845				
RA	Gollnick S.O., Tronstine M.L., Yamashita L.C., Kehry M.				
RT	Moore K.W.				
RT	"Isolation, characterization, and expression of cDNA for				
RT	the mouse Fc receptor for IgE (Fc epsilon R1)."				
RL	J. Immunol. 144:1974-1982(1990).				
RN	[3]				
RP	SEQUENCE FROM N.A. (FORMS 1 AND 2).				
RX	STRAIN=BA/22				
RA	MEDLINE=94372613; PubMed 8066246				
RT	Kondo H., Ichikawa Y., Nakamura K., Isobe Y.S.				
RT	"Cloning of cDNAs for new subtypes of murine low-affinity				
RT	for IgE (Fc epsilon R1/Clz)."				
RL	Int. Arch. Allergy Immunol. 105:48-48(1994).				
RN	[4]				
RP	3D-STRUCTURE MODELING OF THE INTRAMEMBRANE				
RX	MEDLINE=94191542; PubMed=8142907				
RA	Padian E.A., Helm B.A.				
RT	"Modeling of the lectin-homology domains of the human an-				
RT	affinity Fc epsilon receptor (Fc epsilon R1/CD23)."				
RL	Receptor 3:425-441(1993).				
CC	FUNCTION: THIS RECEPTOR HAS ESSENTIAL ROLES IN THE RE-				
CC	COGNITION AND IN THE DIFFERENTIATION OF B CEL-				
CC	LS. A B-CELL-SPECIFIC ANTIGEN.				
CC	-1- SUBCELLULAR LOCATION: Type II membrane protein.				
CC	-1- ALTERNATIVE PRODUCTS: 3 isoforms: A (SHOWN HERE), B				
CC	and C, produced by ALTERNATIVE SPLICING.				
CC	-1- PTM: N- AND O-GLYCOSYLATED (BY SIMILARITY).				
CC	-1- MISCELLANEOUS: THERE ARE TWO KINDS OF FC RECEPTORS				
CC	DIFFER IN BOTH STRUCTURE AND FUNCTION: HIGH AFFINITY				
CC	ISOFORMS AND MAST CELLS AND LOW AFFINITY RECEPTORS				
CC	AND MONOCYTES.				


```

FT REPEAT 111 131
FT DISULFID 160 288 BY SIMILARITY.
FT DISULFID 163 174 BY SIMILARITY.
FT DISULFID 191 282
FT DISULFID 259 273
FT CARBOHYD 63 63 N-LINKED (GLUCNA:...) (POTENTIAL).
FT VARSPIC 1 7 MEGOSYS -> MNPSSQ (IN ISOFORM B).
FT CONFLICT 269 269 N -> T (IN REF. 3).
SU SEQUENCE 421 AA: 36468 MW: 48670.806515887 Cp664:

Query Match: 19.0%, Score 422.5, DB 1, Length 321.
Best Local Similarity: 28.6%; Pred. No. 9e-14;
Matches 95; Conservative 61; Mismatches 121; Indels 55; Gaps 14;

QY 1 MSDEKPRVQGIIGLIGTGHCA--IVGLISFMIAGVIVATIVQSKVPSLSQ--EUSE 57
DQ 1 MEEQVSPFIEELPRRCRCRGTQIVLIGVTAALMAGLLILMLMWDITQSKLDEERA 60
QY 58 QDAIYQMLTQIKAAVG-ELSEKSKLQETVQETQIKAAVEDEPEKSKLQETVQETIKKA 116
DQ 61 ARNVSQVSKNIESHHGDMQAKSQSTQISQLEELKA-----EQQRKKS 104
QY 117 AVGLPEPKSKLQETVQETIKAAVGLPEKSK-----LQETVQETIKKAAGVILPQSK 172
DQ 105 Q--DLELSMNINLQADLSSEKSO--ELNERNESDLEELREPEVTKLR----- 149
QY 173 QQAIVQPIITIKAFERIECHCPKRWTFPQCNVPMNSQPMHDSVACQCPAPALVYI 242
DQ 150 -----MELQVSSGVNTECPQEKMINFCQKCYFGKQTKQVHAKVACQDMDEGLVSI 201
QY 213 FTANEGWETLQTSFSEFTSWMLSTGNEFTWQWVQSTSLSTSTWVWVWVWVWVWVWVW 256
DQ 202 HSEDEEELTRKHSHTD--SWLSTFILFLDEFTWVDSHVYS--NMALDEHSEST 256
QY 292 EECALFESQSG-WNDRGCVQ-NWYICKKPAQ 321
DQ 257 EECVMAKRSQPMNDAPCTPKLCAWCDPIATF 298

RESULT 4
KUCR_RAT STANDARD: PRT: 550 AA.
AC P10716;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE C-type lectin 13 (Kupffer cell receptor).
GN CLEC3L3 OR KCLR.
OS Rattus norvegicus (Prt)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN 11;
RP MEDLINE:88227339; PubMed 2836387.
RA Hoyle G.W., Hill R.L.;
RT "Molecular cloning and sequencing of a cDNA for a carbohydrate
binding receptor unique to rat Kupffer cells."
PI J. Biol. Chem. 263:7487-7492(1998)
RI 12;
RP MEDLINE:91107689; PubMed:1846367;
RA Hoyle G.W., Hill R.L.;
RT "Structure of the gene for a carbohydrate-binding receptor unique to
rat Kupffer cells."
RI J. Biol. Chem. 266:1850-1857(1991)
CC -> FUNCTION: RECEPTOR WITH AN AFFINITY FOR GALACTOSE AND FIUCOSE.
CC COULD BE INVOLVED IN ENOCYTOSIS.
CC -> SUBCELLULAR LOCATION: TYPE II membrane protein.
CC -> TISSUE SPECIFICITY: KUPFER CELLS.
CC -> SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC This SWISS-PROT entry is copyright It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/about
CC or send an email to license@sib-sib.ch).
CC
DR EMBL: J03744, AAA1472.1;
DR EMBL: M6642; AAA0902.1;
DR PIR: A28166; A28166.
DR PIR: A3674; A3674.
DR HSP: P20693; HLU.
DR InterPro: IPR01194, Lectin_C.
DR InterPro: IPR00017; Synactin.
DR Pfam: PF00059; Lectin_C_1.
DR SMART: SM00034; CLECR_1.
DR SMART: SM00504; SYN_1.
DR PROSITE: PS00615; C-TYPE_LECTIN_1;
DR PROSITE: PS50041; C-TYPE_LECTIN_2;
KW Receptor; Transmembrane; Glycoprotein; Lectin; Signal-anchor;
KW Endocytosis.
FT DOMAIN 1 42 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 43 69 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
FT DOMAIN 70 550 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 448 538 C-TYPE LECTIN (SHORT FORM).
FT DISULFID 449 536 BY SIMILARITY.
FT DISULFID 516 528 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLUCNA:...) (POTENTIAL).
FT CARBOHYD 93 93 N-LINKED (GLUCNA:...) (POTENTIAL).
FT CARBOHYD 145 115 N-LINKED (GLUCNA:...) (POTENTIAL).
FT CARBOHYD 142 132 N-LINKED (GLUCNA:...) (POTENTIAL).
FT CARBOHYD 203 229 N-LINKED (GLUCNA:...) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLUCNA:...) (POTENTIAL).
SU SEQUENCE: 550 AA: 61064 MW: 93584.674306270 Cpc64;

Query Match: 18.2%, Score 309.5, DB 1, Length 550;
Best Local Similarity: 28.7%; Pred. No. 1.1e-12;
Matches 81; Conservative 46; Mismatches 110; Indels 45; Gaps 11;

QY 52 SQDESEMAIYQNTQIKAAVELESEKSKLQETVQETQIKAAVGLPEKSK-LQETVQET 110
DQ 286 AQIQNMAHLEFQTTQIQIKAKLKTSTN-----SOLEVANKLKSSPELQTLARD 339
QY 111 LITFAAVTEPEKSKLQETVQETIKAAVSDLEPE-----SKLQETVQETIKK 161
DQ 340 LSVVSA-----KSNVQMSNMQAKAAVQSLKLGELATIKAAVQGSQSDLEALQK 393
QY 163 AAVGFTI PQGSKQGTQVQETITLITAFPELCPIHPKRWTFPQCNVPMNSQPMHDSVAC 221
DQ 394 AVAANHQGTQGVQVQVQIM-----DQWKYFNREYFESDKKSWHEALNF 439
QY 222 QGVRAQIVITKAFPCNPLQIOTSNSPFSWMLSDLSQDQETQWQVQDSELE-ISEGRY 280
DQ 440 QVSGAMLAIVITLQFQAFVLYVLIINAVIQ--WGLDQSTEDENWVQVNIPEVQVSRF 497
QY 281 WNSCHTNG SCN EDCAHFGSGCRGRRCVNWYICK 317
DQ 198 WELSLGEMHIDHELELVVQVQVWWMMAVAVNWCKK 548

RESULT 5
LECH_MOUSE STANDARD: PRT: 283 AA.
AC P34927; Q64363;
DT 01-FEB-1994 (Rel. 28, Created)
DT 15-JUN-1994 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Asialoglycoprotein receptor 1 (hepatic lectin 1) (Mn-1) (ASGP-R)
DE (ASGP-R).
OS Mus musculus (Mouse).
OX NCBI_TaxId=10116;
RN 1;
RP MEDLINE:91107689; PubMed:1846367;
RA Hoyle G.W., Hill R.L.;
RT "Structure of the gene for a carbohydrate-binding receptor unique to
rat Kupffer cells."
RI J. Biol. Chem. 266:1850-1857(1991)
CC -> FUNCTION: RECEPTOR WITH AN AFFINITY FOR GALACTOSE AND FIUCOSE.
CC COULD BE INVOLVED IN ENOCYTOSIS.
CC -> SUBCELLULAR LOCATION: TYPE II membrane protein.
CC -> TISSUE SPECIFICITY: KUPFER CELLS.
CC -> SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC This SWISS-PROT entry is copyright It is produced through a collaboration

```

00 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 01 MRL_TaxID=10090;
 02
 03
 04
 05
 06
 07
 08
 09
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563
 564
 565
 566
 567
 568
 569
 570
 571
 572
 573
 574
 575
 576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609
 610
 611
 612
 613
 614
 615
 616
 617
 618
 619
 620
 621
 622
 623
 624
 625
 626
 627
 628
 629
 630
 631
 632
 633
 634
 635
 636
 637
 638
 639
 640
 641
 642
 643
 644
 645
 646
 647
 648
 649
 650
 651
 652
 653
 654
 655
 656
 657
 658
 659
 660
 661
 662
 663
 664
 665
 666
 667
 668
 669
 670
 671
 672
 673
 674
 675
 676
 677
 678
 679
 680
 681
 682
 683
 684
 685
 686
 687
 688
 689
 690
 691
 692
 693
 694
 695
 696
 697
 698
 699
 700
 701
 702
 703
 704
 705
 706
 707
 708
 709
 710
 711
 712
 713
 714
 715
 716
 717
 718
 719
 720
 721
 722
 723
 724
 725
 726
 727
 728
 729
 730
 731
 732
 733
 734
 735
 736
 737
 738
 739
 740
 741
 742
 743
 744
 745
 746
 747
 748
 749
 750
 751
 752
 753
 754
 755
 756
 757
 758
 759
 760
 761
 762
 763
 764
 765
 766
 767
 768
 769
 770
 771
 772
 773
 774
 775
 776
 777
 778
 779
 780
 781
 782
 783
 784
 785
 786
 787
 788
 789
 790
 791
 792
 793
 794
 795
 796
 797
 798
 799
 800
 801
 802
 803
 804
 805
 806
 807
 808
 809
 810
 811
 812
 813
 814
 815
 816
 817
 818
 819
 820
 821
 822
 823
 824
 825
 826
 827
 828
 829
 830
 831
 832
 833
 834
 835
 836
 837
 838
 839
 840
 841
 842
 843
 844
 845
 846
 847
 848
 849
 850
 851
 852
 853
 854
 855
 856
 857
 858
 859
 860
 861
 862
 863
 864
 865
 866
 867
 868
 869
 870
 871
 872
 873
 874
 875
 876
 877
 878
 879
 880
 881
 882
 883
 884
 885
 886
 887
 888
 889
 890
 891
 892
 893
 894
 895
 896
 897
 898
 899
 900
 901
 902
 903
 904
 905
 906
 907
 908
 909
 910
 911
 912
 913
 914
 915
 916
 917
 918
 919
 920
 921
 922
 923
 924
 925
 926
 927
 928
 929
 930
 931
 932
 933
 934
 935
 936
 937
 938
 939
 940
 941
 942
 943
 944
 945
 946
 947
 948
 949
 950
 951
 952
 953
 954
 955
 956
 957
 958
 959
 960
 961
 962
 963
 964
 965
 966
 967
 968
 969
 970
 971
 972
 973
 974
 975
 976
 977
 978
 979
 980
 981
 982
 983
 984
 985
 986
 987
 988
 989
 990
 991
 992
 993
 994
 995
 996
 997
 998
 999
 1000

00 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 01 MRL_TaxID=10090;
 02
 03
 04
 05
 06
 07
 08
 09
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 5

KW Receptor, Transmembrane; Glycoprotein; Lectin; Signal-anchor;
 Endocytosis.
 FT DOMAIN 1 42 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 43 69 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
 FT DOMAIN 70 548 EXTRACELLULAR (POTENTIAL)
 FT DOMAIN 438 538 C-TYPE LECTIN (SHORT FORM).
 FT DISULFID 440 536 BY SIMILARITY.
 FT DISULFID 516 528 BY SIMILARITY.
 FT CARBOHYD 86 86 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 92 92 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 115 115 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 132 132 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 209 209 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 255 255 N-LINKED (GLCNAC) (POTENTIAL).
 SQ SOURCE 548 AA: 61268 MW: 6764951820E73BD9 CRR64.
 Query Match: 16.8%; Score 284.5; DB 1; Length 548;
 Best Local Similarity: 26.0%; Pred No 4 1e-11;
 Matches: 80; Conservative: 58; Mismatches: 115; Indels: 55; Gaps: 13;
 QY 40 ILVGVSVPSLSQSEQDAIYQNTOLKA---AVGEISE-KSKLQETVQL----- 88
 ID 256 ISALQIVPGQMPKPKMSLELEETLTAQYKANGHLPQDQVJOLKALEETGSL 217
 QY 89 -TOLKAAVRLPKSK-TQETVQELTLKAAVGELEPKSLQETVQELTLKAAVGELE 146
 DB 116 NSRIEVNQCKMDASRELOILKRLSLVSAL-----KSNVQMLQSNQRAETNGTL- 467
 QY 147 KSKQ-----ETVQELTL-----KAAVGELEPKSQOQIYQELTDLKTAEELCRHCP 195
 DB 168 KADLDATKAI TAIQYQEQN/ALQFVAAQKQFQKQVQNCVQLQLA----- 413
 QY 196 KDWTFQGCYFNSQSPNHNDSVTAQGVFACI VIKTAEFCNFIQLOTSNPFSSMC 255
 DB 414 QNMKYENGNYEYSBOKPWRBAEKCTSGAGLASVTSQEDQALVQTTSSGDH--WIG 471
 QY 256 LSLNCTGTQWYTGSLSTSEFP-YVNSSEIR-----NSCHELAEISSGSNNLAF 109
 DB 472 LLDGCTGTPWVGTIPFNNAQSKGFWKNNVIMPHRNSFEPLVAVH-VLQNNLMAVSS 540
 QY 410 DNYWICKK 317
 DB 531 SYRWCKK 538
 RESULT 10
 LECH HUMAN
 ID LECH HUMAN STANDARD; PRT, 290 AA.
 AC P07306;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Asialoglycoprotein receptor 1 (Hepatic lectin II) (ASGPR) (ASGP-R).
 GN ASGRI.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Molesteostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN 11
 RP SOURCE FROM N.A.
 RX MEDLINE=8513041; PubMed=2982798;
 RI Spiess M., Schwartz A.L., Lodish H.F.;
 PI "Sequence of human asialoglycoprotein receptor cDNA. An internal
 PI signal sequence for membrane receptor".
 RL J. Biol. Chem. 260:1979-1982(1985).
 RN 12
 RP SOURCE FROM N.A.
 RX MEDLINE=86079574; PubMed=3753585;
 RI Spiess M., Lodish H.F.;
 PI "An internal signal sequence: the asialoglycoprotein receptor membrane
 PI anchor".
 RL Cell 44:177-185(1986)

RN 13
 RP SOURCE FROM N.A.
 RA Wang H., Guo X., Li L., Lou H., Huang Y., Han J.;
 RI "Human asialoglycoprotein receptor 1 gene is expressed in SH-SY5Y
 RI human neuroblastoma cells".
 PL Submitted (Sep-2001) to the EMBL/GenBank/DBJ databases.
 CC 1- FUNCTION: MEDIATES THE ENDOCYTOSIS OF PLASMA GLYCOPROTEINS TO
 CC WHICH THE TERMINAL SIALIC ACID RESIDE ON THEIR COMPLEX
 CC CARBOHYDRATE MOETIES HAS BEEN REMOVED. THE RECEPTOR RECOGNIZES
 CC TERMINAL GALACTOSE AND N-ACETYLGLUCOSAMINE UNITS. AFTER LIGAND
 CC BINDING TO THE RECEPTOR, THE RESULTING COMPLEX IS INTERNALIZED AND
 CC DEGRADATED. THE RECEPTOR THEN RETURNS TO THE CELL MEMBRANE
 CC SURFACE.
 CC 1- SUBCELLULAR LOCATION: Type II membrane protein.
 CC 1- TISSUE SPECIFICITY: Expressed exclusively in hepatic parenchymal
 CC CELLS.
 CC 1- PTM: PHOSPHORYLATED ON A CYTOPLASMIC SER RESIDUE.
 CC 1- MISCELLANEOUS: CALCIUM IS REQUIRED FOR LIGAND BINDING.
 CC 1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL/GenBank/DBJ
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: M10068; AA051785.1;
 CC EMBL: AB070943; BAB83508.1;
 CC DR PIR: A22509; LNH01.
 CC DR HSSP: P06744; LHL1.
 CC DR Genbank: BC02742; ASGRI.
 CC MIM: 108460;
 CC DR InterPro: IP001304; Lectin_C.
 CC Pfam: PF00559; Lectin_C1.
 CC SMART: SM00044; CLEST_1.
 CC PROSITE: PS00427; 2-TYPE_LECTIN_1.
 CC PROSITE: PS00441; 2-TYPE_LECTIN_2_1.
 CC FEELSLER: PS00441; 2-TYPE_LECTIN_2_1.
 KW Lectin; Glycoprotein; Receptor; Endocytosis; Transmembrane;
 KW Calcium; Signal-anchor; Phosphorylation.
 FT INIT_MET 0 0
 FT DOMAIN 1 39
 FT TRANSMEM 40 60
 FT TRANSMEM 40 60
 FT DOMAIN 61 290
 FT DOMAIN 152 278
 FT SILE 4 7
 FT DISULFID 153 164 BY SIMILARITY.
 FT DISULFID 181 276 BY SIMILARITY.
 FT DISULFID 254 268 BY SIMILARITY.
 FT CARBOHYD 80 80 N-LINKED.
 FT CARBOHYD 148 148 O-LINKED.
 SQ SOURCE 290 AA: 34054 MW: 818970E30DAE1586 CRR64.
 Query Match: 16.0%; Score 271.5; DB 1; Length 290;
 Best Local Similarity: 27.0%; Pred. No. 1.3e-10;
 Matches: 86; Conservative: 45; Mismatches: 112; Indels: 75; Gaps: 13;
 QY 6 EPVVOGILALYGLGCAVMIQILSPMLAGVVAIIYOVSKVPSLSQSEQDAIYQNT 65
 DB 28 QPILQRI-----GSGPKILSLD-----GLSLDLVVVGVGSONSLOEIRKIRKIF 76
 QY 66 LQAAVGLSEKSLAEYVQIIGAAVSIPEKSKGLHYELIPLKAAVCHLITERS 125
 DB 77 SNFIASH-----EAGVKGLISLGGGWVGGKMSLE-----S 106
 QY 126 KIGFVGHITLTKAAVGLPLKSKLQETVQELTLKAAVGELEPKSLQETVQELTLKT 185
 DB 107 GPEKQGMDSH-----QNSN-----LHVKQVPSIDLSQMAAQ-----GN 146
 QY 186 AFERVTHETPEATFTQCNVYKQNGERHMEISVIAQGVNAGLVITAEQONTAIGT 245

```

147 GSEKTC--CPNWWVEHESCTWFSSTKAMAMANNVPIETAHLLVAVTSWEECKEYVHHI 204
246 SSKSNF-SWNGSLDNOESTWQVDSPLSPSFQYWNSSGFPNN-----STNEI*AFPSG 269
205 GGVN -TWGGLHD--QNGFKNWVGFVETFTFRN-WFFEPFLWY-HGTFGEI*AFFTD 289
400 SG-WNDKGDVDNWK 316
268 WSPMNDVQRPYRWCE 277

RESULT 11
LECI_HUMAN
ID LECI_HUMAN STANDARD: PRT: 311 AA.
AC P07307: 000448: Q04969:
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Asialoglycoprotein receptor 2 (hepatic lectin H2) (ASGP-R) (ASGPR).
EN ASGPR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mollusca;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
GX NCBI_TaxID=9606;
LN 11
SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=46016723; PubMed=1463106;
RA Spiess M., Lodish H.F.;
RT "Sequence of a second human asialoglycoprotein receptor: conservation of two receptor genes during evolution."
RI Proc. Natl. Acad. Sci. U.S.A. 82:6465-6469(1985).
RN 12
SEQUENCE FROM N.A. (ISOFORMS 2 AND 3).
KC TISSUE=Liver;
KA MPRLNE-92184202; PubMed 1377982;
RA Paletta E., Stockert R.J., Racevskis J.;
RT "Differences in the abundance of variably spliced transcripts for the second asialoglycoprotein receptor polypeptide, H2, in normal and transformed human liver."
RI Hepatology 15:395-402(1992).
RN 14
SEQUENCE FROM N.A. (ISOFORM 3).
KC TISSUE=Liver;
RA Strausberg R.;
RI Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases
CC -1- FUNCTION: MEDIATES THE ENDOCYTOSIS OF PLASMA GLYCOPROTEINS TO WHICH THE TERMINAL SIALIC ACID RESIDUE ON THEIR COMPLEX CARBOHYDRATE MOETIES HAS BEEN REMOVED. THE RECEPTOR RECOGNIZES TERMINAL GALACTOSE AND N-ACETYLGLUCOSAMINE UNITS. AFTER LIGAND BINDING TO THE RECEPTOR, THE RESULTING COMPLEX IS INTERNALIZED AND TRANSPORTED TO A SORTING ORGANELLE, WHERE RECEPTOR AND LIGAND ARE DISASSOCIATED. THE RECEPTOR THEN RETURNS TO THE CELL MEMBRANE SURFACE.
CC -1- SUBUNIT: THE FUNCTIONING LIGAND-BINDING UNIT OF THIS RECEPTOR IS THOUGHT TO BE AT LEAST A DIMER.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein.
CC -1- ALTERNATIVE PRODUCTS: 3 isoforms: 1 (shown here), 2 and 3; are produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN HEPATIC PARENCHYMAL CELLS.
CC -1- MISCELLANEOUS: CALCIUM IS REQUIRED FOR LIGAND BINDING.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL, contribution of the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
CC EMBL: M11025; AAB59519.1;

```

```

DR EMBL: U97197; AAB58308.1;
DR EMBL: X55283; CAA38997.1;
DR EMBL: AC017251; AAH17251.1;
DR FTR: A5179; LNHQ2A.
DR HSSP: P06734; 1HL1.
DR GCGW: HGNC:743; ASGPR2.
DR MIM: 108361;
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C_1.
DR SMART: SMC0034; CLACT_1.
DR PROSITE: PS00615; C-TYPE_LECTIN_1;
DR PROSITE: PS50041; C-TYPE_LECTIN_2;
KW Lectin; Glycoprotein Receptor; Endocytosis; Transmembrane;
KW Calcium; Signal anchor; Phosphorylation; Alternative splicing; GLYCOPOLASMIC (POTENTIAL);
FT TRANSMEM 5 79
FT DOMAIN 80 311
FT DOMAIN 176 402
FT SITE 5 8
FT D1SULEFD 177 188
FT D1SULEFD 205 400
FT D1SULEFD 278 292
FT CARBOHYD 192 102
FT CARBOHYD 170 170
FT CARBOHYD 305 305
FT MOD_RES 12 12
FT VARSPIC 24 42
FT VARSPIC 82 86
FT SEQUENCE 311 AA; 35191 MW; 840784FCFEBA16 CRC64.

Query Match: 15.7%; Score 267; DR 1; Length 311
Best Local Similarity: 27.2%; Ident. No. 276-10; Indels 0
Matches 89; Conservative 49; Mismatches 103; Indels 0

UY 6 EPRVQALALALALGALVALALSLMLAVLVALIVVSVKPSLSLEQSI
LB 47 QLAARLRLSMVFS-----LALSLNHL--LTVICV-----
UY 66 TQKAAVCHLSPKSKLOEYLOHLOKAAVCHLPEKSKLOEYLOHLOKAA
DB 77 -----TGSQSEGHGALQALQALPRLSKAEFSNF--SSHLTEV--QALSTHQS
UY 126 KQEIYQFETLKAAVCHLPEKSKLOEYLOHLOKAAVCHLPEKSKLOEYLOHLOKAA
DB 124 -----KITSLGA-----KILKQGLKAKHDAHLFLKHPVCLPH
UY 178 QELDLKTAFLERLCRHCPEKWTFTQGNFYKNSQSRWHSVTAQGEVRAAL
DB 168 H-----SMSSQRTG--CPVNWVEHESCTWFSSTKAMAMANNVPIETAHLLVAVTSWEECKEYVHHI
UY 248 CNFLOLOLSKRSKRSKRWMSLSLNOBGTWQVDSPLSPSFQYWNSSGFPNN
DB 221 GKEIVQHNPN--TWGLID--SNSSKAWVADYDKHYNKN--WAVTLPNNA
UY 292 EPCAFESFSG--WNHFCDEVCHWYLYK 317
DB 276 ECGVEVGFQDGKWNHDFCLQVTRWYWK 402

RESULT 12
LECI_MOUSE
ID LECI_MOUSE STANDARD: PRT: 301 AA.
AC P24721:
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Asialoglycoprotein receptor 2 (hepatic lectin 2) (MHL-2) (ASGPR).
EN ASGPR2 OR ASGP-R.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

```



```

RI  Megabalanus rosa ";
RL  Biochim Biophys. Acta 874:285-295(1986).
CC  -1- FUNCTION: SUGAR-BINDING PROTEIN WHICH PROCOGNIZES SPECIFIC
CC  CARBOHYDRATE STRUCTURES AND AGGLUTINATE A VARIETY OF ANIMAL
CC  CELLS BY BINDING TO CELL-SURFACE GLYCOPROTEINS AND GLYCOLIPIDS.
CC  THIS IS A CALCIUM-DEPENDENT LECTIN. INVERTEBRATE LECTINS MAY BE
CC  INVOLVED IN DEFENSE FUNCTIONS.
CC  -1- SUBUNIT: HOMOTETRAMER; DISULFIDE-LINKED.
CC  -1- TISSUE SPECIFICITY: COELEMIC FLUID.
CC  -1- MISCELLANEOUS: THIS LECTIN BINDS GALACTOSE.
CC  -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/uniformer/
CC  or send an email to license@sib-sib.ch).
CC  -----
DR  EMBL: D13302; HAA02556.1; -
DR  EMBL: D13299; HAA02556.1; JOINED.
DR  EMBL: D13300; HAA02556.1; JOINED.
DR  EMBL: D13301; HAA02556.1; JOINED.
DR  PIR: A26094; ENRC3.
DR  HSSP: P23806; 1XX.
DR  InterPro: IPR001304; LECTIN_C.
DR  Pfam: PF00059; LECTIN_C.1.
DR  SMART: SM00034; CLECT.1.
DR  PROSITE: PS00615; C-TYPE_LECTIN_1; 1.
DR  PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
KW  Lectin; calcium; signal.
FT  SIGNAL      1      24
FT  CHAIN       1      162
FT  DOMAIN      25      152
FT  DISULFID    26      39
FT  DISULFID    56      150
FT  DISULFID    125     142
FT  DISULFID    157     157
FT  DISULFID    160     160
FT  VARIANT     146     146
FT  VARIANT     146     146
SO  SOURCE: 162 AA; 18328 MW; BR7F14FQ1D01GB1 CPG64.

Query Match      12.9%; Score 219.5; DB 1; Length 162;
Host Local Similarity 32.5%; Pred. No. 1.3e-07;
Matches 51; Conservative 30; Mismatches 61; Indels 15; Gaps 7;

QY  172 KNGQIVYETTRIKTAFEPLCPHP-----KQWTFEGQNTYFMSNSQPNMHDSTAFQYR- 226
DB  2 QRSFIVQAVTLVVVAITTAETCPGNDLWQEDYDHCYASTYQVRWMDAQLAQVHP 61

QY  227 -AQLVVIKTAEEONFLOLOTSRSNRPMMGI SPLNDEGTWQVDSPLSPSPQRYWNSGE 285
DB  62 GAYLATIQSOLENAFLS-ETVSNRRL-WIGLNDIDLEGHYVWSNGE--ATDF-TYWSSNN 116

QY  286 PANS:NFT-A-----EFSGSAMNLNFTVFNWLCRR 318
DB  117 PNMWENQDCGVVYDVTGQMDDDCNKNKNFLCKMP 153

```

Search completed: December 7, 2002, 11:02:24
 Job time : 30 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2002, 11:06:55, Search time: 34 seconds

(without alignments)
1273.719 Million cell updates/sec

Title: US-09-831-458A-12

Perfect score: 325
Sequence: 1 MSLSKRPVQQLGLQVIGH PRTVINWICKPAACEPDE 325

Scoring table:
Gapop 60.0, Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size: 5

Total number of hits satisfying chosen parameters: 27600

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: A-Geneseq_101002.*

```

1: /SID52/gcdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcdata/geneseq/geneseq-emb1/AA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	325	100.0	325	21	AAV94345
2	168	51.7	296	20	AAW88493
3	168	51.7	296	22	AAH79093
4	168	51.7	296	23	AAH09715
5	97	29.8	216	22	AAH94720
6	78	24.0	404	14	AAH32188
7	78	24.0	404	21	AAH19714
8	78	24.0	404	21	AAH28614
9	78	24.0	404	22	AAH79086
10	78	24.0	404	23	AAH80025

11	55	16.9	144	22	AAH55444
12	50	15.4	50	22	AAH42034
13	50	15.4	50	22	AAH25647
14	50	15.4	50	22	AAH62914
15	50	15.4	50	22	AAH75726
16	50	15.4	50	22	AAH55875
17	50	15.4	50	23	AAH645211
18	5	2.8	438	22	AAV72948
19	9	2.8	394	18	AAH25742
20	9	2.8	751	18	AAH01825
21	9	2.8	751	23	AAH14366
22	9	2.8	778	22	AAH58942
23	9	2.8	811	23	AAH14381
24	9	2.8	879	23	AAH74756
25	9	2.8	918	22	AAH50945
26	9	2.8	918	23	AAH24351
27	9	2.8	918	23	AAH19481
28	9	2.8	926	23	AAH24350
29	9	2.8	935	23	AAH24353
30	9	2.8	963	23	AAH24352
31	8	2.5	36	22	AAH72965
32	8	2.5	36	22	AAH72966
33	8	2.5	111	22	AAH72964
34	8	2.5	212	21	AAH23492
35	8	2.5	213	21	AAH24291
36	8	2.5	239	22	AAH78158
37	8	2.5	279	21	AAH26290
38	8	2.5	325	23	AAH88024
39	8	2.5	325	23	AAH88026
40	8	2.5	356	23	AAH25448
41	8	2.5	472	23	AAH81511
42	8	2.5	500	22	AAH95815
43	8	2.5	764	20	AAH24094
44	7	2.2	27	21	AAH45133
45	7	2.2	29	12	AAH22106

ALIGNMENTS

RESULT 1	AAV94345
AAV94345	standard: Protein: 325 AA.
XX	22-A05-2000 (first entry)
XX	Human cell surface receptor protein #12.
XX	Human: HCSR, cytosolic; antiarthritic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neutrophilic; eosinophilic; antineoplastic; cancer; lung melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy.
XX	Human sapiens.
XX	Homo sapiens.
XX	Location/Qualifiers
XX	1..52
XX	/label: Signal_peptide
XX	53..325
XX	/label: HCSR-12
XX	46..44
XX	/label: Transmembrane_domain
XX	1..48
XX	/note: "2" fully immunoglobulin receptor"
XX	22..69
XX	/note: "2" fully immunoglobulin receptor"
XX	95..142
XX	/note: "2" fully immunoglobulin receptor"
XX	118..165
XX	/note: "2" fully immunoglobulin receptor"

```

FI Region 56 /note= "potential phosphorylation site"
FI Region 64 /note= "potential glycosylation site"
FI Region 76 /note= "potential phosphorylation site"
FI Region 185 /note= "potential phosphorylation site"
FI Region 211 /note= "potential phosphorylation site"
FI Region 234 /note= "potential phosphorylation site"
FI Region 245 /note= "potential phosphorylation site"
FI Region 248 /note= "potential phosphorylation site"
FI Region 287 /note= "potential phosphorylation site"
FI Region 289 /note= "potential glycosylation site"
FI Region /note= "potential phosphorylation site"
PN W0200028032-A2.
XX
XX 18-MAY-2000.
XX
XX 12-NOV-1999; 99WO-US26742.
XX
XX 12-NOV-1998; 98US-0191280.
XX 07-DEC-1998; 98US-0206647.
XX 08-MAR-1999; 99US-0123404.
XX
XX (INCYTE PHARM INC.
PI Tang YT, Corley NC, Guegler KJ, Yue H, Baughn MR, Lal P;
PI Hillman JL, Bandman O, Azimzai Y, Au-Young J;
PI WPI: 2000-376546/32.
XX
XX New human cell surface receptor protein and polynucleotide useful for
PI diagnosis, prevention and treatment of cancer, immune disorders,
PI infection and neuronal disorders.
PI
PI Claim 1: Page 86-87; 97pp; English.
XX
XX The present sequence is a novel human cell surface receptor protein
XX (HCSR) designated HCSR-12. The nucleotide sequence was identified in
XX cDNA clone 3344986 from the cDNA library SPLN0109, which was made
XX from RNA isolated from diseased spleen tissue. A number of Incyte Clones
XX were used to assemble the consensus sequence. BLAST analysis showed that
XX the sequence is homologous to non-CD4 glycoprotein gp120 receptor
XX AAR42188. HCSR and its antagonist are useful for preventing or treating
XX disorders associated with decreased or increased expression or activity
XX of HCSR. Such disorders include cancers such as leukemia and melanoma,
XX immune disorders such as rheumatoid arthritis, asthma and
XX atherosclerosis, bacterial and parasitic infections and neuronal
XX disorders such as akathisia, Alzheimer's disease, multiple sclerosis and
XX epilepsy. Polynucleotides encoding HSRs may be used as hybridisation
XX probes to diagnose these conditions. Anti-HCSR antibodies may be used
XX as antagonists, as a targeting or delivery mechanism for binding
XX pharmaceutical agents into contact with cells or tissues expressing
XX HCSR and for diagnosis of HCSR-related disorders. HCSR and its
XX catalytic or immunogenic fragments are useful for drug screening using
XX libraries of compounds.
XX
XX Sequence 325 AA:
XX
XX Query Match 100.0%; Score 225; DB 21; Length 325.
XX Best Local Similarity 100.0%; From No. 4 2e-244;
XX Matches 32/32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 MSNSKRRVVOGLIIGTIGGALVIGIISPMILAGVIVATVQVSKVPSLSIQPSQPDAA 60
XX 41 MSNSKRRVVOGLIIGTIGGALVIGIISPMILAGVIVATVQVSKVPSLSIQPSQPDAA 60
XX

```

```

XX 61 TVNLTQGLAAVTELSKSKLQETVQVETQIKAAVDELPSKSLQETVQETPLKAAVGE 120
XX 61 TVNLTQGLAAVTELSKSKLQETVQVETQIKAAVDELPSKSLQETVQETPLKAAVGE 120
XX 121 EPRSKSLQETVQETPLKAAVDELPSKSLQETVQETPLKAAVDELPSKSLQETVQET 180
XX 121 EPRSKSLQETVQETPLKAAVDELPSKSLQETVQETPLKAAVDELPSKSLQETVQET 180
XX 121 EPRSKSLQETVQETPLKAAVDELPSKSLQETVQETPLKAAVDELPSKSLQETVQET 180
XX 181 EPRSKSLQETVQETPLKAAVDELPSKSLQETVQETPLKAAVDELPSKSLQETVQET 240
XX 181 EPRSKSLQETVQETPLKAAVDELPSKSLQETVQETPLKAAVDELPSKSLQETVQET 240
XX 241 EPRSKSLQETVQETPLKAAVDELPSKSLQETVQETPLKAAVDELPSKSLQETVQET 300
XX 241 EPRSKSLQETVQETPLKAAVDELPSKSLQETVQETPLKAAVDELPSKSLQETVQET 300
XX 301 GMDNPEVDPNVEIKKPACTHDE 325
XX 301 GMDNPEVDPNVEIKKPACTHDE 325
XX
XX RESULT 2
XX AAW88493
XX ID AAW88493 standard; Protein; 296 AA.
XX
XX AC AAW88493;
XX
XX 10-MAR-1999 (first entry)
XX
XX Human liver clone H01347-encoded transmembrane protein.
XX
XX Transmembrane protein; H01347; human; lectin; receptor; liver.
XX
XX Homo sapiens.
XX
XX W0985508 A2.
XX
XX 10-DEC-1998.
XX
XX 03-JUN-1998; 98WO-JP02445.
XX
XX 03-JUN-1997; 97JP-0144948.
XX
XX (PROT-) PROTEOME INC.
XX (SAGA) SAGAMI CHEM RES CENTRE.
XX
XX Kato S, Sekine S, Yamaguchi T;
XX
XX WPI: 1999-045740/04.
XX N-PSDB: AAW84361.
XX
XX New human proteins containing transmembrane domains and their
XX encoding sequences useful in the preparation of antibodies and
XX large-scale protein production, gene diagnosis, and gene therapy
XX
XX Claim 1: page 125-126; 18pp; English.
XX
XX This is the amino acid sequence of a transmembrane protein encoded
XX human liver cDNA clone H01347 (see AAW84361). The encoded protein
XX is characterised as a type-II membrane protein having a single
XX N-terminal transmembrane domain. It has 85.6% homology over 284
XX amino acid residues to the human HIV envelope glycoprotein gp120-
XX binding C type lectin (a CD4 independent HIV receptor) and may
XX function as a receptor on the membrane surface. The invention
XX provides nucleotide sequences (see AAW84359-76) coding for 18
XX transmembrane proteins (see AAW88491-508), vectors containing such
XX polynucleotides, and eukaryotic cells containing the vectors. The
XX proteins can be used as antigens or as compositions in the
XX preparation of antibodies against the proteins. The polynucleotides
XX can be used as probes for gene diagnosis, and as gene sources for
XX gene therapy and large-scale production of proteins encoded by the
XX cDNA. The host cells are used for the detection of ligands

```


CC corresponding to the expressed proteins, and the screening of low
 CC mol.wt. medicines.

XX Sequence 296 AA;

Query Match 51.7%, Score 168, ID 23, Length 296;

Best Local Similarity 100.0%; Pred. No. 1,2e-150;

Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSDSKEPRVQGLGILGCHGALVGLLSPMLAGVVALVAVSVSPSSLSQESQSDA 69

Db 1 MSDSKEPRVQGLGILGCHGALVGLLSPMLAGVVALVAVSVSPSSLSQESQSDA 60

QY 61 IYQNTQLKAAGVGLSEKSKLQEIYQELTQLKAAGVGLPEKSKLQEIYQELTRKAAGVF 120

Db 61 IYQNTQLKAAGVGLSEKSKLQEIYQELTQLKAAGVGLPEKSKLQEIYQELTRKAAGVF 120

QY 121 LPEKSKLQEIYQELTRKAAGVGLPEKSKLQEIYQELTRKAAGVGLP 168

Db 121 LPEKSKLQEIYQELTRKAAGVGLPEKSKLQEIYQELTRKAAGVGLP 168

RESULT 4

ID AAG79093 standard; Protein: 296 AA

XX AAG79093;

XX 10-DEC-2001 (first entry)

XX Amino acid sequence of a splice variant of human DC-SIGN.

XX Human receptor; DC-SIGN; dendritic cell; T lymphocyte; HIV;

XX gp120; C-type lectin; ICAM3; HIV entry; T cell; macrophage;

XX HIV infection; splice variant.

XX Homo sapiens.

XX W0200164752-A2.

XX 07-SEP-2001.

XX 28-FEB-2001; 2001W0-US06322.

XX 02-MAR-2000; 2000US-0517605.

XX (UNYV) UNIV NEW YORK STATE.

XX (UNYV-) UNIV NJMCCEN.

XX Liltman DK, Kwon D, Van Kooyk Y, Geljtenbreck T;

XX WPI: 2001-602565/6A.

XX N-PSDB: AAI65469.

XX An antibody for the treatment or prevention of HIV infection comprises

XX a gp120 portion which binds to DC-SIGN or is exposed upon gp120 binding

XX of DC-SIGN due to concomitant conformational change.

XX Example 1: Page 130-131; 131pp; English.

XX The present sequence represents a splice variant of a human receptor

XX designated DC-SIGN. DC-SIGN is specifically expressed on dendritic cells

XX and facilitates infection of T lymphocytes with HIV. DC-SIGN is identical

XX to a HIV-1 gp120-binding C-type lectin. DC-SIGN binds ICAM-3 (which is

XX expressed constitutively on T lymphocytes) with high affinity. The

XX specificity describes an antibody which is specific for an antigenic

XX fragment of gp120. This antigenic fragment binds to DC-SIGN or is exposed

XX upon gp120 binding or DC-SIGN due to concomitant conformational change.

XX The antibody inhibits the trans enhancement of HIV entry into a T

XX cell or macrophage facilitated by dendritic cells. The antibody is

XX useful to treat or prevent HIV infection.

XX Sequence 296 AA;

Query Match 51.7%, Score 168, ID 23, Length

Best Local Similarity 100.0%; Pred. No. 1,2e-150;

Matches 168; Conservative 0; Mismatches 0; Indels 0;

QY 1 MSDSKEPRVQGLGILGCHGALVGLLSPMLAGVVALVAVSVSPSSLSQESQSDA 69

Db 1 MSDSKEPRVQGLGILGCHGALVGLLSPMLAGVVALVAVSVSPSSLSQESQSDA 60

QY 61 IYQNTQLKAAGVGLSEKSKLQEIYQELTQLKAAGVGLPEKSKLQEIYQELTRKAAGVF 120

Db 61 IYQNTQLKAAGVGLSEKSKLQEIYQELTQLKAAGVGLPEKSKLQEIYQELTRKAAGVF 120

QY 121 LPEKSKLQEIYQELTRKAAGVGLPEKSKLQEIYQELTRKAAGVGLP 168

Db 121 LPEKSKLQEIYQELTRKAAGVGLPEKSKLQEIYQELTRKAAGVGLP 168

RESULT 4

ID ABB09715 standard; Protein: 296 AA.

XX ABB09715;

XX 11-JUN-2002 (first entry)

XX Amino acid sequence of human polypeptide HP01347.

XX Antibody; antigen; transmembrane domain protein; HP01347

XX Homo sapiens.

XX W0200208416-A1.

XX 31-JAN-2002.

XX 24-JUL-2001; 2001W0-JP06371.

XX 24-DEC-2000; 2000JP-0222743.

XX 24-AUG-2000; 2000JP-0254407.

XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX Kato S, Nagata N, Fujimura N, Kobayashi M, Ito K

XX WPI: 2002-195877/25.

XX N-PSDB: ABL41993.

XX Antibody preparation by inoculation of an animal with a

XX transmembrane domain for use as drugs, diagnostic reagents

XX laboratory reagents.

XX Example; Page 23-27; 45pp; Japanese.

XX The specification describes a method of antibody preparation

XX method comprises inoculating an animal with a vector carrying

XX a fusion protein having an antigen protein fused to the

XX side (extracellular) of a transmembrane domain protein

XX N-terminal side of which is intracellular, and then

XX purifying the antibody from the animal. The antibodies

XX as drugs, vaccines, diagnostic reagents and laboratory

XX present sequence represents a polypeptide designated by

XX used in the course of the invention.

XX Sequence 296 AA;

Query Match 51.7%, Score 168, ID 23, Length

Best Local Similarity 100.0%; Pred. No. 1,2e-150;

Matches 168; Conservative 0; Mismatches 0; Indels 0;

QY 1 MSDSKEPRVQGLGILGCHGALVGLLSPMLAGVVALVAVSVSPSSLSQESQSDA 69

Db 1 MSDSKPRVQGLTGLACTGHALVIGLISPMILAVLVALVGVKSPSSI SQSSEQDA 60
 QY 61 IYQMLTQKAAVGLSKSLQGLTYOELTQKAAVGLPEKSKLOETTYOELTRKAAVGL 120
 Db 61 IYQMLTQKAAVGLSKSLQGLTYOELTQKAAVGLPEKSKLOETTYOELTRKAAVGL 120
 QY 121 LPEKSKLOETTYOELTRKAAVGLPEKSKLOETTYOELTRKAAVGL 168
 Db 121 LPEKSKLOETTYOELTRKAAVGLPEKSKLOETTYOELTRKAAVGL 168
 RESULT 5
 AAB94720
 ID AAB94720 standard; Protein: 216 AA.
 AC AAB94720;
 XX 26-JUN-2001 (first entry)
 DE Human protein sequence SEQ ID NO:15736.
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 OS Homo sapiens.
 FN EPI074617-A2.
 ID 07-FEB-2001.
 XX 28-JUL-2000; 2000EP-0116126.
 XX 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118774.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX (HELI-) HELIX RES INST.
 FA Oda T, Iisoda T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DE WPI: 2001-318749/34.
 XX
 PR Primer sets for synthesizing polynucleotides, particularly the 5602
 full-length cDNAs defined in the specification, and for the detection
 and/or diagnosis of the abnormality of the proteins encoded by the
 full-length cDNAs -
 PS Claim 8; SEQ ID 15736; 2537pp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 full-length cDNAs defined in the specification. Where a primer set
 comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 to the complementary strand of a polynucleotide which comprises one of
 the 5602 nucleotide sequences defined in the specification, where the
 oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 of an oligonucleotide comprising a sequence complementary to the
 complementary strand of a polynucleotide which comprises a 5'-end
 sequence and an oligonucleotide comprising a sequence complementary to a
 polynucleotide which comprises a 3'-end sequence, where the
 oligonucleotide comprises at least 15 nucleotides and the combination of
 the 5'-end sequence/3'-end sequence is selected from those defined in
 the specification. The primer sets can be used in antisense therapy and
 in gene therapy. The primers are useful for synthesizing polynucleotides,
 particularly full-length cDNAs. The primers are also useful for the
 detection and/or diagnosis of the abnormality of the proteins encoded by
 the full-length cDNAs. The primers allow obtaining of the full-length
 cDNAs easily without any special methods. AAB94720 to AAB17432 and
 AAB17433 to AAB18742 represent human cDNA sequences; AAB92446 to
 AAB95894 represent human amino acid sequences; and AAB13629 to AAB17432
 represent oligonucleotides, all of which are used in the exemplification
 of the present invention.

XX
 SQ Sequence 216 AA:
 Query Match 29.8%; Score 97; DB 22; Length 216;
 Best local similarity 100.0%; Pred. No. 1,46-83;
 Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 72 VQPLSEKSKLOETTYOELTQKAAVGLPEKSKLOETTYOELTRKAAVGLPEKSKLOETTY 131
 Db 44 VQPLSEKSKLOETTYOELTQKAAVGLPEKSKLOETTYOELTRKAAVGLPEKSKLOETTY 103
 QY 132 QELTRKAAVGLPEKSKLOETTYOELTRKAAVGL 168
 Db 104 QELTRKAAVGLPEKSKLOETTYOELTRKAAVGL 140
 RESULT 6
 AAR32188
 ID AAR32188 standard; Protein: 404 AA.
 AC AAR32188;
 XX 08-JUN-1993 (first entry)
 DE Sequence of a non CD4 glycoprotein gp120 receptor protein.
 KW Surface cell protein, glycoprotein, receptor protein, gp120 receptor,
 HIV; diagnosis; therapy.
 OS Homo sapiens.
 FN W03301820-A.
 FN W03301820-A.
 FD 04-FEB-1993.
 XX 16-JUL-1992; 92MO-US05985.
 XX 16-JUL-1991; 91US-0731214.
 XX (BRIM) BRISTOL-MYERS SQUIBB CO.
 PI Curtis RM;
 DE WPI: 1993-058515/07.
 DR N-PSDB; AAQ36648.
 XX Inhibition of non-CD4 mediated HIV infection dependent on novel
 gp-120 receptor, for diagnosing and treating HIV infection
 PS Disclosure; Figure 3A; 44pp; English.
 XX
 CC The gp120 receptor protein is present on placental, muscle, neural,
 brain, dendritic or mucosal cells. It may be produced recombinantly
 by cloning cDNA isolated from a library of recombinant placental
 genes. Binding of gp120 to its receptor is inhibited by specific
 carbohydrates, plant lectins and specific antibodies. It corresp.
 to a non-CD4 mammalian cell surface protein, contg. ca. 400 AAs and
 having a mol. wt. of ca. 15 kD, and a binding affinity for gp-120 of
 kd = 1.32-2 nM.
 SQ Sequence 404 AA:
 Query Match 24.0%; Score 76; DB 14; Length 404;
 Best local similarity 100.0%; Pred. No. 2,56-65;
 Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 72 GPATVGLTQKAAVGLPEKSKLOETTYOELTRKAAVGLPEKSKLOETTYOELTRKAA 117
 Db 74 GPATVGLTQKAAVGLPEKSKLOETTYOELTRKAAVGLPEKSKLOETTYOELTRKAA 133
 QY 118 VQPLSEKSKLOETTYOELT 135
 Db 134 VQPLSEKSKLOETTYOELT 151

RESULT 7
 AAB19714
 ID AAB19714 standard; Protein: 404 AA.
 XX
 AC AAB19714;
 XX
 DT 19-FEB-2001 (first entry)
 XX
 DE Dendritic cell specific C-type lectin DC-SIGN.
 XX
 KW Dendritic cell; lectin; DC-SIGN; intercellular adhesion molecule;
 KM ICAM-3 receptor; human; HIV-1; infection; autoimmune disease;
 XX allergy; immunotherapy; immunosuppressive; therapy.
 XX
 OS Homo sapiens.
 XX
 PN EPI046651-A1.
 XX
 PD 25-OCT-2000.
 XX
 PP 19-APR-1999; 99EP-0201204.
 XX
 PR 19-APR-1999; 99EP-0201204.
 XX
 PA (UYN1-) UNIV KONINK NIDMEGEN.
 XX
 PI Fijodor CG, Geijlenbeek TH, Van Kooyk Y, Torensmo P;
 XX
 DR WP1: 2000-048929/63.
 XX
 DR N-PSDB: AAB88740.
 XX
 PT Modulating the immune response in an animal, useful e.g. for
 PT immunosuppression, using a compound that binds to C-type lectin on
 PT dendritic cells -
 XX
 PS Claim 11: Page 18-20: 44pp, English.

CC The present sequence is that of human DC-SIGN, a 44 kDa C-type
 CC lectin of dendritic cells (DC). DC-SIGN is a novel ICAM-3 receptor
 CC that is specifically expressed by human DC and is involved in the
 CC initial transient DC-T-cell interaction necessary for initiating an
 CC immune response. DC-SIGN is also able to bind to the HIV envelope
 CC protein gp120 and to facilitate HIV-1 entry into DC. The invention
 CC relates to the use of a compound that binds to a C-type lectin
 CC (especially DC-SIGN) on the surface of a DC, and in the preparation
 CC of a composition for modulating, especially reducing, the immune
 CC response in an animal, in particular a human or other mammal. The
 CC composition modulates the interactions between a DC and a T-cell,
 CC more specifically between a C-type lectin on the surface of a DC
 CC and an ICAM receptor on the surface of a T-cell. It is used to
 CC prevent/inhibit immune responses to specific antigens, for inducing
 CC tolerance, for immunotherapy, for immunosuppression, for the
 CC treatment of autoimmune diseases, the treatment of allergy, and/or
 CC for inhibiting HIV infection (claimed). The compound that binds to
 CC the C-type lectin is chosen from mannose, fucose, plant lectins,
 CC antibiotics, sugars, proteins and antibodies against C-type
 CC lectins, especially against DC-SIGN. Anti-DC-SIGN antibodies were
 CC developed that not only inhibited transient DC-T-cell interactions
 CC and DC-induced T-cell proliferation but also inhibited HIV-1
 CC infection of DC.
 XX
 XX
 SU Sequence 404 AA:
 XX
 Query Match 24.0%; Score 78; DB 21; Length 404;
 Best Local Similarity 100.0%; Pred. No. 2,5e-65;
 Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QDAIYONLTOLKAAVGLSEKSLQEIYQELTOLKAAVGLPEKSKLQEIYQELTILKAA 117
 Db 74 QDAIYONLTOLKAAVGLSEKSLQEIYQELTOLKAAVGLPEKSKLQEIYQELTILKAA 133

QY 118 VGLPEKSKLQEIYQELT 135
 Db 134 VGLPEKSKLQEIYQELT 151

RESULT 8
 AAB28614
 ID AAB28614 standard; Protein: 404 AA.
 XX
 AC AAB28614;
 XX
 DT 12-FEB-2001 (first entry)
 XX
 DE Human C-type lectin receptor.
 XX
 KW Human; immunomodulatory; antifungal; anti-HIV; anti-
 KW antidiabetic; antithyroid; antihemorrhagic; antiarthritic;
 KW C-type lectin receptor; immunotherapy; immunosuppressive;
 KW (transplant rejection); autoimmune disease; thyroiditis;
 KW rheumatoid arthritis; multiple sclerosis; autoimmune dis-
 KW systemic lupus erythematosus; HIV infection; allergy.
 XX
 OS Homo sapiens.
 XX
 PN W0200063251-A1.
 XX
 PD 26-OCT-2000.
 XX
 PP 19 APR 2000; 2000W0 NL02555.
 XX
 PR 19-APR-1999; 99EP-0201204.
 XX
 PR 20-JAN-2000; 2000CS-0174224.
 XX
 PA (UYN1-) UNIV NIDMEGEN.
 XX
 PI Fijodor CG, Geijlenbeek TH, Van Kooyk Y, Torensmo P;
 XX
 DR WP1: 2000-058424/63.
 XX
 DR N-PSDB: AAC65383.
 XX
 PT Methods for modulating dendritic cell and T cell inter-
 PT action, such as mannose carbohydrates, that binds to
 PT on a dendritic cell surface, useful in the treatment of
 PT diseases and allergy -
 XX
 PS Disclosure: Fig 9: 66pp; English.

CC The present sequence is for SIGN, a 44kDa C-type lectin;
 CC methods for modulating dendritic cell and T cell inter-
 CC action, such as mannose carbohydrates, that binds to
 CC disclosed. The methods involve using a compound that binds
 CC lectin on the surface of a dendritic cell. The methods
 CC preventing or inhibiting immune responses to specific an-
 CC tigen, for immunotherapy, for immunosuppression, for the
 CC preventing transplant rejection), for the treatment of
 CC diseases (e.g. thyroiditis, rheumatoid arthritis, multi-
 CC autoimmune diabetes, systemic lupus erythematosus), HIV
 CC and/or for the treatment of allergy. Antibodies are used
 CC detecting the presence of dendritic cells in a biological
 CC for determining the presence and/or expression of the
 CC or their fragments or epitopes in a biological sample.
 XX
 XX
 SU Sequence 404 AA:
 XX
 Query Match 24.0%; Score 78; DB 21; Length 404;
 Best Local Similarity 100.0%; Pred. No. 2,5e-65;
 Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QDAIYONLTOLKAAVGLSEKSLQEIYQELTOLKAAVGLPEKSKLQEIYQELTILKAA 117
 Db 74 QDAIYONLTOLKAAVGLSEKSLQEIYQELTOLKAAVGLPEKSKLQEIYQELTILKAA 133

1b 134 VGELEPKSKLOEIYQELT 151

RESU1.9
AAG79086 standard: Protein: 404 AA.
XX
XX AAG79086;
AC
XX 10-DEC-2001 (first entry)
DE Human DC-SIGN, a dendritic cell-specific C-type lectin.
XX
XX
XX Human receptor DC-SIGN, dendritic cell; T lymphocyte; HIV;
KM gp120; C type lectin; ICAM3; HIV entry; T cell; macrophage;
KM HIV infection.
XX
XX Homo sapiens.
XX
XX W0200164752-A2.
XX
XX 07-SEP-2001.
XX
XX 28-FEB-2001; 2001WO-0536722
XX
XX 02-MAR-2000; 2000US-0517605
XX
XX (UYNY) UNIV NEW YORK STATE.
XX (UYNY-) UNIV NITMGEN
XX
XX Litman DP, Kwag D, Van Kowyk Y, Gelfenbeek T;
DR WPI: 2001-602565/68
DR N-PSDB: AA165461.
XX
XX An antibody for the treatment or prevention of HIV-infection comprises
PT a gp120 portion which binds to DC-SIGN or is exposed upon HIV-infection
PT of DC-SIGN due to concomitant conformational change.
XX
XX Example 1: Page 114-115, 131pp; English.
XX
XX The present sequence represents a human receptor designated DC-SIGN, that
CC is specifically expressed on dendritic cells and facilitates infection
CC of T lymphocytes with HIV. DC-SIGN is identical to a HIV-1
CC gp120 binding C-type lectin. DC-SIGN binds ICAM 3 (which is expressed
CC constitutively on T lymphocytes) with high affinity. The specification
CC describes an antibody which is specific for an antigenic fragment of
CC gp120. This antigenic fragment binds to DC-SIGN or is exposed upon
CC gp120 binding of DC-SIGN due to concomitant conformational change.
CC The antibody inhibits the trans enhancement of HIV entry into a T
CC cell or macrophage facilitated by dendritic cells. The antibody is
CC useful to treat or prevent HIV infection
XX
XX Sequence 404 AA:
SQ

Query Match 24.0%; Score 78; DB 23; Length 404;
Best local similarity 100.0%; Pred. No. 2, 5e-65;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 58 GDAIVYNTLQKAAGVSEKSKLOEIYQELTQKAAGVLEPKSKLOEIYQELTRKAA 117
DB 74 GDAIVYNTLQKAAGVSEKSKLOEIYQELTQKAAGVLEPKSKLOEIYQELTRKAA 133
UY 118 VGELEPKSKLOEIYQELT 135
DB 134 VGELEPKSKLOEIYQELT 151

RESULT 10
AA08R025
ID AA08R025 standard: Protein: 404 AA.
XX
XX AA08R025;
AC

XX 05-JUN-2002 (first entry)
DT
XX
XX Mouse C1S1-B7 ORF protein.
DE
XX
XX Nucleic acid library: immune response; asthma; COPD;
KM airway hyperresponsiveness; bronchoalveolar macrophage;
KM signature sequence; SS; chronic obstructive pulmonary disease;
KM allergic disease; rhinitis; atopic dermatitis; urticaria;
KM autoimmune disease; multiple sclerosis; inflammatory bowel disease;
KM allograft rejection; infectious disease;
KM calcium activated chloride channel.
XX
XX Mus sp.
XX
XX W020021466-A2.
XX
XX 21-FEB-2002.
XX
XX 16-APR-2001; 2001WO-NL00610.
XX
XX 16-APR-2000; 2000EP-0202867.
XX
XX (UYNY) UNIV SUNNY DRECHT.
XX
XX Grodz PC, Van Bergenhenegouwen BJ, Van Oosterhout AJM;
DR WPI: 2002-241886/29.
XX
XX Nucleic acid library comprising genes which are capable of initiation,
PT progression and suppression of an immune response, especially an immune-
PT response observed with airway hyper-responsiveness of asthma.
XX
XX Example 10; Fig 11; 120pp; English.
XX
XX The invention relates to a nucleic acid library comprising genes or
CC their fragments which are capable of modulating an immune response
CC observed with airway hyper-responsiveness and/or bronchoalveolar
CC manifestations of asthma. Also included are a method for modulating an
CC immune response of an individual comprising modulating a gene comprising
CC a nucleic acid at least functionally equivalent to a nucleic acid
CC identifiable by SS for the production of an antagonist (for use as a
CC substance (for use as a modulator) capable of modulating a gene
CC comprising a nucleic acid at least functionally equivalent to a nucleic
CC acid identifiable by SS and the use of a prophylactic substance derived
CC from a nucleic acid at least functionally equivalent to a nucleic acid
CC identifiable by SS for the production of an antagonist (for use as a
CC medicament) against the substance. The antagonist and substance are
CC useful for the treatment of an immune response observed with airway
CC hyper-responsiveness and/or bronchoalveolar manifestations of asthma
CC The method is useful for modulating the above immune response, where the
CC gene encodes a gene product capable of modulating the immune response
CC The substance is useful for treating an immune response, particularly
CC asthma, chronic obstructive pulmonary disease (COPD), allergic diseases
CC (rhinitis, atopic dermatitis, urticaria), autoimmune diseases (e.g.
CC multiple sclerosis), inflammatory bowel disease, allograft rejection and
CC infectious disease. The present sequence is a mouse or human
CC protein encoded by a signature sequence gene or its homologue/functionally
CC equivalent.
XX
XX Sequence 404 AA:
SQ

Query Match 24.0%; Score 78; DB 23; Length 404;
Best local similarity 100.0%; Pred. No. 2, 5e-65;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 58 GDAIVYNTLQKAAGVSEKSKLOEIYQELTQKAAGVLEPKSKLOEIYQELTRKAA 117
DB 74 GDAIVYNTLQKAAGVSEKSKLOEIYQELTQKAAGVLEPKSKLOEIYQELTRKAA 133
UY 118 VGELEPKSKLOEIYQELT 135
DB 134 VGELEPKSKLOEIYQELT 151

Db 144 VSELPKSKTQETVQELT 151

RESULT 11

AA025424 standard: Protein: 143 AA.

AA025424:

16-OCT-2001 (first entry)

Human protein sequence SEQ ID NO:939.

Human, cancer, ulcer, HIV infection, human immunodeficiency virus; antiinflammatory; antihemorrhagic; antidiabetic; immunosuppressive; antibacterial; endocrine; cardiac; central nervous system; virucide; anti-HIV; fungicide; antimutagen; cardiovascular; antineoplastic; antiaggregant; haemostatic; vulnery; antitumor; osteopathic; eczema; dermatological; antiallergic; antistimulant; antidiabetic; cytostatic; neuroprotective; antidepressant; footrotic; antiparkinsonian; infection; immunostimulant; gene therapy; antisense therapy; vaccine; inflammation; antidiarrheal; rheumatoid arthritis; septic shock; pancreatitis; cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmune; genetic disease; haematopoietic disorder; platelet disorder; asthma; thrombocytopenia; osteoporosis; severe combined immunodeficiency; allergic rhinitis; diabetes; multiple sclerosis; depression; Alzheimer's disease; Parkinson's disease; neurodegenerative disorder; neurological disorder.

Homo sapiens.

W0200153455-A2.

26-JUL-2001.

22-DEC-2000; 2000W0-0835017.

23-DEC-1999; 99US-0471275.

21-JAN-2000; 2000US-0488725.

25-APR-2000; 2000US-0752477.

(HYSE-) HYSE0 INC.

Tang YT, Liu C, Drmanac RT;

WPI: 2001-457603/49

N-PSDB: AA099365.

Claim 20; Page 202; 1217pp; English.

AA099166 to AA099904 encode the human proteins given in AA025225 to AA025563. The proteins can have activities based on the tissues and cells they are expressed in, such as: antidiarrheal; antihemorrhagic; antiallergic; immunosuppressive; antibacterial; endocrine; cardiac; central nervous system; virucide; anti-HIV; fungicide; antimutagen; cardiovascular; antineoplastic; antiaggregant; haemostatic; vulnery; antitumor; osteopathic; dermatological; antiallergic; antistimulant; antidiabetic; cytostatic; neuroprotective; antidepressant; footrotic; antiparkinsonian; infection; immunostimulant; gene therapy; antisense therapy and vaccine production. The proteins and polynucleotides are useful for screening for agonists or antagonists of a protein and for the treatment and diagnosis of disorders associated with the activity of a protein e.g. inflammation, rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction, neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal infections, autoimmunity, genetic diseases, haematopoietic disorders, anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers, osteoporosis, severe combined immunodeficiency, eczema, allergic rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression, Alzheimer's disease, Parkinson's disease, neurodegenerative and

CC neurological disorders.

XX Sequence 143 AA:

Query Match 16.94; Score 55; EB 22; Length

Best Local Similarity 100.00; Prod. No. 5,60-44;

Matches 55; Conservative 0; Mismatches 0; Trunc:

0Y 58 QDAIYONLTQKAAVSELSKSLQETVQELTQKAAVSELPKSKTQELT

Db 72 QDAIYONLTQKAAVSELSKSLQETVQELTQKAAVSELPKSKTQELT

RESULT 12

AB042033 standard: Peptide: 50 AA.

AB042033:

04-FEB-2002 (first entry)

Peptide #9539 encoded by human foetal liver single exon

Human, foetal liver, gene expression; single exon nucle

Homo sapiens.

W0200157277-A2.

09-AUG-2001.

30-JAN-2001; 2001W0-US00669.

04-FEB-2000; 2000US-0180412.

26-MAY-2000; 2000US-0207456.

18-JUN-2000; 2000US-0608408.

03-AUG-2000; 2000US-0632366.

21-SEP-2000; 2000US-0244687.

27-SEP-2000; 2000US-0246359.

04-OCT-2000; 2000US-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Kank DK;

WPI: 2001 483447/52.

Human genome-derived single exon nucleic acid probes use

analyzing gene expression in human foetal liver -

Claim 27; SEQ ID NO 34668; 679pp - sequence listing: 80

XX Sequence 50 AA:

Query Match 15.48; Score 50; DB 22; Length

Best Local Similarity 100.00; Prod. No. 1,20-39;

Matches 50; Conservative 0; Mismatches 0; Trunc:

0Y 189 KLRHCEPKDWIFFQGNLYMSNSQRMHDSVTAQVEPAQIVIKTAPEL

Db 1 KLRHCEPKDWIFFQGNLYMSNSQRMHDSVTAQVEPAQIVIKTAPEL

```

RESULT 14
AM625647
ID ABR25647 standard; Protein: 50 AA.
XX
AM625647:
XX
ID 23-JAN-2002 (first entry)
XX
DE Protein #7646 encoded by probe for measuring heart cell gene expression.
XX
XX Human; gene expression; heart; microarray; vascular system;
XX cardiovascular disease; hypertension; cardiac arrhythmia;
XX congenital heart disease.
XX
OS Homo sapiens.
XX
PN M0200157274-A2.
XX
ID 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-0506647.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0638408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX
PS Claim 15, SEQ ID NO 27417, 530pp, English
XX
CC The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart (see
XX ABA1533-AAA11303). The present sequence is a protein encoded by one such
XX probe. The probes may be used for predicting, monitoring and diagnosing
XX gene expression in samples derived from the human heart via microarrays.
XX By measuring gene expression, the probes are useful for predicting,
XX diagnosing, grading, staging, monitoring and prognosing diseases of the
XX human heart and vascular system e.g. cardiovascular disease,
XX hypertension, cardiac arrhythmias and congenital heart disease.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at http://wipo.int/pub/published\_pat\_sequences.
XX
SQ Sequence 50 AA:
XX
Query Match 15.4% Score 50; DB 23; Length 50;
Best local Similarity 100.0%; Pred. No. 1,2e-19;
Matches 50, Conservative 0, Mismatches 0, Indels 6, Gaps 0,
XX
CY 189 FTTTFTTFTTFTTTCGTCYFMSNSQPMNHISVTAAGPFAQIVTKTAFQ 238
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
ID 1 RUCRHTNLTWFTTTCGTCYFMSNSQPMNHISVTAAGPFAQIVTKTAFQ 50
XX
RESULT 14
AAM62914
ID AAM62914 standard; Protein: 50 AA.
XX
AAM62914:
XX
XX 05-NOV-2001 (first entry)
XX

```

```

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 35019.
XX
XX Human; brain expressed exon; gene expression analysis; probe;
XX microarray; Alzheimer's disease; atherosclerosis; schizophrenia;
XX epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN M0200157275-A2.
XX
ID 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-0506647.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0638408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX
PS Example 4; SEQ ID NO: 35019; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is a protein encoded by one of
XX the probes of the invention.
XX
SQ Sequence 50 AA:
XX
Query Match 15.4% Score 50; DB 23; Length 50;
Best local Similarity 100.0%; Pred. No. 1,2e-19;
Matches 50, Conservative 0, Mismatches 0, Indels 6, Gaps 0,
XX
CY 189 FTTTFTTFTTFTTTCGTCYFMSNSQPMNHISVTAAGPFAQIVTKTAFQ 238
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
ID 1 FTTTFTTFTTFTTTCGTCYFMSNSQPMNHISVTAAGPFAQIVTKTAFQ 50
XX
RESULT 15
AAM75726
ID AAM75726 standard; Protein: 50 AA.
XX
AAM75726:
XX
XX 06-NOV-2001 (first entry)
XX
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 36032.
XX Human; brain expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN M0200157276-A2.
XX
ID 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-0506648.
XX

```

```

PE 04-FEB-2000; 2000US-0180312
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366
PR 21-SEP-2000; 2000US-02346A7
PR 27-SEP-2000; 2000US-0246359
PR 04-OCT-2000; 2000CB-0024263.
XX
FA (MOLE-) MOLECULAR DYNAMICS INC.
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488900/53.
XX
PS Example 4: SEQ ID NO: 36032, 658bp + Sequence Listing; English.
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
SO Sequence 50 AA;
XX
Query Match 15.4%; Score 50; DI 22; Length 50;
Post Local Similarity 100.0%; Pred. No. 1.2e-39;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 189 RLCRHCRCXKDWTFPGGCMCYFMSNSGRMHHSVATCGFPACIIVIKIAEEQ 238
Db 1 RLCRHCRCXKDWTFPGGCMCYFMSNSGRMHHSVATCGFPACIIVIKIAEEQ 50
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
RESULT 16
ID AAM35835 standard; Protein: 50 AA.
AA AAM35835;
AC AAM35835;
DT 17-OCT-2001 (first entry)
DE Peptide #9872 encoded by probe for measuring placental gene expression.
KW Prob; microarray; human; placenta; antenatal] diagnosis;
KM genetic disorder.
XX Homo sapiens.
XX W0200157272-A2.
PN XX
PD 09-AUG-2001.
PE 30-JAN-2001; 2001WO-0S90663
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-02346A7.
PR 27-SEP-2000; 2000US-0246359.
PR 04-OCT-2000; 2000CB-0024263.
XX
FA (MOLE-) MOLECULAR DYNAMICS INC.
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-48897/53.
XX
Human genome-derived single exon nucleic acid probes useful for

```

PT	analyzing gene expression in human placenta	
XX		
PS	Claim 27; SEQ ID No 36104; 644pp; English.	
XX		
CC	The present invention relates to single exon nucleic acid	
CC	see A413115-A4157546). The present sequence is a peptide	
CC	such probe. The probes are useful for producing a micro	
CC	predicting, measuring and displaying gene expression in	
CC	from human placenta. The probes are useful for antenata	
CC	human genetic disorders.	
XX		
SO	Sequence 50 AA:	
UY	Query Match: 15.4%; Score 50; DB 22; Length:	
UB	Best Local Similarity: 100.0%; Prod. No. 1-2e-39;	
	Matches 50; Conservative 0; Mismatches 0; Indels	
UY	189 RLCRCPCPKDWTFFQGCNCYEMSNLSJRNHDSYACQVRAQIVIKIAERQ	
UB	1 RLCRCPCPKDWTFFQGCNCYEMSNLSJRNHDSYACQVRAQIVIKIAERQ	
RESULT 17		
ID	ABG45211	
AC	ABG45211 standard; Peptide: 50 AA.	
XX		
XX	ABG45211;	
XX		
DT	19 AUG:2002 (first entry)	
DE	Human peptide encoded by genome-derived single exon pro	
XX		
XX	Human: single exon probe; asthma; lung cancer; COPD; l	
KM	chronic obstructive pulmonary disease; interstitial lung	
KM	lamellar idiopathic pulmonary fibrosis; neurofibromatosis	
KM	tuberosus sclerosis; Gaucher's disease; Niemann-Pick disc	
KM	Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemo	
KM	primary histiocytosis; lymphangioleiomyomatosis; Kartag	
KM	primary alveolar proteinosis; fibrocystic pulmonary d	
KM	primary ciliary dyskinesia; pulmonary hypertension;	
KM	hyaline membrane disease.	
XX		
OS	Homo sapiens.	
PN	WO200186903-A2.	
XX		
PD	15-NOV-2001.	
XX		
PE	30-JAN-2001; 2001WO-US006545.	
XX		
PR	04-FER-2000; 2000US-180312P.	
PR	26-MAY-2000; 2000US-207456P.	
PR	30-JUN-2000; 2000US-060840P.	
PF	03-AUG-2000; 2000US-061336P.	
PF	21-SEP-2000; 2000US-244527P.	
PE	27-SEP-2000; 2000US-246359P.	
PE	04-OCT-2000; 2000GB-002426A.	
XX		
PA	(MOLE-) MOLECULAR DYNAMICS INT.	
XX		
PI	Peim SC; Hanzel DK; Chen W; Rank BR;	
XX		
UK	WPI: 2002-114183/15.	
XX		
FT	Spatially addressable set of single exon nucleic acid	
PT	measure gene expression in human lung samples	
XX		
PS	Claim 27; SEQ ID No 34876; 644pp; English.	
XX		
CC	The invention relates to a spatially-addressable set of	
CC	nucleic acid probes for measuring gene expression in a	
CC	from human lung comprising single exon nucleic acid pro	
CC	12614 nucleic acid sequences mentioned in the specificat	

complements of the 12387 open reading frames derived from the 12614 probes. Also included are a microarray comprising the novel set of probes; the novel set of probes which hybridise at high stringency to a nucleic acid expressed in the human lung; measuring gene expression in a sample derived from human lung, comprising (a) contacting the array with a collection of detectably labeled nucleic acids derived from human lung mRNA, and (b) measuring the label detectably bound to each probe of the array; identifying exons in a eukaryotic genome, comprising (a) algorithmically predicting at least one exon from genomic sequences of the eukaryote, and (b) detecting specific hybridisation of detectably labeled nucleic acids from eukaryote lung mRNA, to a single exon probe, having a fragment identical to the predicted exon, the probe is included in the above mentioned microarray; assigning exons to a single gene, comprising (a) identifying exons from genomic sequence by the method above, and (b) measuring the expression of each of the exons in several tissues and/or cell types using hybridisation to a single exon microarray having a probe with the exon, where a common pattern of expression of the exons in the tissues and/or cell types indicates that the exons should be assigned to a single gene; a peptide comprising one of 12011 sequences, mentioned in the specification, or encoded by the probes/open reading frames (ORF). The probes are used for gene expression analysis, and for identifying exons in a gene, particularly using human lung derived mRNA and for the study of lung diseases such as asthma, lung cancer, chronic obstructive pulmonary disease (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Huntington's disease, Friedreich's ataxia, haemochromatosis, pulmonary histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis, Karagen syndrome, fibrocystic pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension and hyaline membrane disease. The present sequence is a peptide/protein encoded by a single exon probe of the invention.

Note: The sequence data for this patent did not form part of the prior art specification, but was obtained in electronic format directly from Wipo at http://wipo.int/pub/published_pcl_sequences.

Sequence 50 AA:
 Query Match 15.4%; Score 50; DR 23; Length 50;
 Best Local Similarity 100.0%; Pct. No. 1 2e-39;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 189 RCRHRCPRKMTFFGNCYFMSNSGRNHHVSTACGEVAVQVLVTKIAEQ 238
 1 RCRHRCPRKMTFFGNCYFMSNSGRNHHVSTACGEVAVQVLVTKIAEQ 50

RESULT 18
 AAY72948
 ID AAY72948 standard; Protein: 238 AA.

AC AAY72948;
 DT 13-JUN-2001 (first entry)

DE Mouse type II integral membrane protein, CIRE.

KW Mouse; type II integral membrane protein; CIRE; anti-allergic vaccine; immunosuppressive; autoimmune disorder; allergy; viral infection; APC; antigen presenting cell; bacterial infection; macrophage; dendritic cell; etc.

OS Mus musculus.

PN W020011986-9-A1.

PD 22-MAR-2001.

PF 13-SEP-2000; 2000WO-AU01096.

PR 13-SEP-1995; 99AU-0002788.

PA (COEN) CHANG-CHI EBERHARDT INST MEDICAL RES
 PI Caminschi I, Wraith MD, Shortman KD;
 DR WPI: 2601-244786/25.
 DR N-PSDB: AM094221.

XX Novel type II integral membrane protein, CIRE, preferentially expressed in dendritic cells, macrophages and their precursors stimulatory to T cells; useful for identifying compounds having immunoregulatory activity.

Claim 1: Fig 1; 53pp; English.

The present sequence is mouse type II integral membrane protein (designated as CIRE) which is preferentially expressed in dendritic cells, macrophages and their precursors stimulatory to T cells. CIRE protein has a N-terminal cytoplasmic domain, transmembrane domain and a C-type lectin domain in the extracellular region. CIRE is useful for screening immunoregulatory compounds and as a marker for dendritic cell (DC) population. CIRE specific ligand such as monoclonal antibodies are useful for isolating an antigen presenting cell (APC) from a biological sample. CIRE DNA is also used as tools to analyse the properties and functions of the CIRE gene/protein. For example, the DNA molecules are used to generate animal models, which lack functional CIRE genes and to isolate regulatory regions of the CIRE gene. Such regulatory regions are used to selectively express exogenous genes in DC or APC. Ligands are useful for modulating immune responses by interacting with the function, migration or maturation of DC or APC. The ligands bind to and interfere with the function of myeloid dendritic cells such that antigen processing is undertaken by lymphoid DC which leads to immune suppression and allergy for treating allergies and autoimmune disorders. The ligands are also useful to block or enhance the interaction of viruses or bacteria with DC, for improving resistance to infection. These ligands are also used to target molecules such as vaccine components, to DC or APC.

Sequence 238 AA:
 Query Match 2.8%; Score 9; DR 22; Length 238;
 Best Local Similarity 100.0%; Pct. No. 3.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 85 YQHTQTKA 93
 1 YQHTQTKA 99

RESULT 19
 AAM25722
 ID AAM25722 standard; Protein: 394 AA.

AC AAM25722;
 DT 08-MAY-1998 (first entry)

DE Human partial beta meltrin protein fragment 2.

KW Meltrin; myotube formation; muscle tissue; therapy; membrane protein; myoblast; adhesion; fusion; medicinal; muscle disorder; human.

OS Homo sapiens.

PN W09731109-A1.

PD 28-AUG-1997.

PF 17-OCT-1996; 96WO-JP03017.

PR 23-FEB-1996; 96JP-0061756.

PA (MORI) MOCHIDA PHARM CO LTD.

PR Fujisawa A, Mizushima O, Okawa N, Shirakawa K, Yamakawa T;


```

XX  WP: 1997-435161/40.
DR  N-PSDB: AAT86014
XX  Meltrin proteins involved with myotube formation in muscle - contain
PT  metalloproteinase and disintegrin domains
XX
PS  Claim 5: Fig 17: 262pp; Japanese.
XX
CC  AAW25721 and AAW25722 represent fragments of a new human membrane
CC  protein, beta meltrin, which participates in the adhesion and fusion of
CC  myoblasts in the course of myotube formation in muscle tissue. Meltrins
CC  and their partial peptides can be used in the study of myotube formation,
CC  and in medicinal compositions for muscle disorders. Meltrins contain an
CC  intracellular domain, a membrane-penetrating domain, a fusion peptide
CC  domain, a disintegrin domain, a metalloproteinase domain, a precursor
CC  domain and a cysteine rich domain although these are not shown in the
CC  specification. Meltrin antagonists can be used as bone absorption
CC  inhibitors, e.g. for the treatment of osteoporosis and hypercalcaemia
CC  and for prevention of cancer cell infiltration.
XX
SU  Sequence 394 AA:
XX
Query Match 2.88; Score 9; DB 18; Length 394;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
UY 34 AGVLVAILV 42
DB 333 AGVLVAILV 341

```

RESULT 20

AAW01825
ID AAW01825 standard; Protein; 751 AA.

XX AAW01825;
XX
XX 27-MAY-1997 (first entry)
XX
XX Human metalloproteinase.
XX
XX Metalloproteinase; infection; inflammation; allergy; therapy.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Domain 26..240
XX /label= Catalytic_domain
XX
XX W09709420-A2.
XX
XX 13-MAR-1997.
XX
XX 05-SEP-1996; 96MO-GH02180.
XX
XX 11-JUN-1996; 96GB-0012145.
XX 05-SEP-1995; 95GB-0018026.
XX 20-OCT-1995; 95GB-0021496.
XX 20-OCT-1995; 95GB-0021497.
XX 21-DEC-1995; 95GB-0026230.
XX
XX (GILT) CELLTech THERAPEUTICS LTD.
XX
XX Dochterly ALP, Slocombe PM;
XX
XX WP: 1997-192896/17.
XX N-PSDB: AAT62559.
XX
XX DNA encoding human metalloproteinase used to develop prods. for
XX treating, e.g. infectious diseases, inflammatory disease or allergy
XX
XX Claim 5; page 5; 17pp; English.

```

XX  A novel human metalloproteinase (AAW01825), or its cDNA,
CC  can be expressed in prokaryotic or eukaryotic host cell.
CC  an isolated gene sequence (see also AAT62558-59). The
CC  metalloproteinase can be used to screen for cpds. that
CC  its activity, and to raise antibodies. Such cpds. and
CC  are useful for the in vivo treatment of diseases or dis-
CC  related to over or under regulated prodn. of the enzyme
XX
SU  Sequence 751 AA:
XX
Query Match 2.88; Score 9; DB 18; Length 751
Best Local Similarity 100.0%; Pred. No. 9; 1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
UY 34 AGVLVAILV 42
DB 530 AGVLVAILV 538

```

RESULT 21

AAO14366
ID AAO14366 standard; Protein; 751 AA.

XX AAO14366;
XX
XX 24-MAY-2002 (first entry)
XX
XX Human metalloproteinase-like/disintegrin-like protein #.
XX
XX Human; metalloproteinase-like/disintegrin-like protein; #.
XX
XX angiogenesis; endothelial cell migration; integrin ligand;
XX
XX ocular disorder; cancer; inflammation; osteoporosis; osteo-
XX
XX cystostatic; antiinflammatory; osteopathic; vasotropic; pro-
XX
XX vulnervary; antiallergic; antibacterial; virucide; proin-
XX
XX antinaeemic; analgesic; thrombosis; tissue repair; neuro-
XX
XX infection.
XX
XX Homo sapiens.
XX
XX W0200210406-A2.
XX
XX 07-FEB-2002.
XX
XX 27-JUL-2001; 2001W0-US23734.
XX
XX 28-JUL-2000; 2000US 221848P.
XX
XX (IMMUNEX CORP.
XX
XX Dubose RF, Wiley SP, Black KA;
XX
XX WP: 2002-195965/25.
XX
XX Novel purified human metalloproteinase-disintegrin poly-
XX
XX for treating ocular and inflammatory disorders; osteopor-
XX
XX testosis, thrombosis, and chronic pain conditions
XX
XX Claim 1, Fig 1; 101pp; English.
XX
XX The present invention provides a number of human metallo-
XX
XX /disintegrin-like (MP) proteins and their fragments. The
XX
XX to identify treatments for ocular disorders, malignant
XX
XX conditions, inflammatory diseases, osteoporosis and other
XX
XX mediated by accelerated bone resorption, restenosis, throm-
XX
XX platelet activation, recruitment or aggregation, thrombo-
XX
XX requiring tissue repair or wound healing, endothelial mal-
XX
XX angiogenesis, allergies, reproductive, neurological and
XX
XX conditions, infections, cardiovascular disorders such as
XX
XX infarction, chronic pain conditions, endocrine system dis-
XX
XX gastrointestinal system disorders, genitourinary system
XX
XX anaemia and haematological disorders. The present sequen-
XX
XX of the invention.

RESULT 23
AA014381
ID AA014381 standard; Protein; A11 AA
XX ,

KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;

XX WP: 2002-479761/51.
 DR N-PSDB: AAB9150.
 XX
 PT New nucleic acid encoding a protein that shares sequence similarity
 PT with mammalian meltrin beta/ADAM 19 homology metalloproteinase for use in
 PT diagnosis and treatment, and in drug screening -
 XX
 PS Claim 1: Page 38-40; 48pp; English.
 XX
 CC The invention relates to novel human proteins (NHP), novel human
 CC proteases that shares sequence similarity with mammalian meltrin beta/
 CC ADAM 19 homologue metalloproteinase, and their corresponding nucleic acid
 CC sequences. NHP DNA is useful for diagnosis, drug screening, clinical
 CC trial monitoring, treatment of diseases and disorders and
 CC pharmacological, cosmetic and nutritional applications. It is also
 CC useful as hybridisation probes for screening libraries, and assessing
 CC gene expression patterns (particularly using a microarray or high-
 CC throughput chip format). It is also useful in restriction fragment
 CC length polymorphism (RFLP) analysis to identify specific individuals,
 CC and as a probe to identify the corresponding mutant allele in the
 CC libraries. It is also useful for the detection of mutant
 CC NHPs or inappropriately expressed NHPs for the diagnosis of disease.
 CC It is also useful in gene therapy. The present amino acid sequence
 CC is novel human protease. Human Nhp gene is located at chromosome 5.
 CC
 SQ Sequence 918 AA:
 Query Match: 2.8%, Score 9, EB 23, Length 918,
 Best Local Similarity 100.0%, Pred. No. 11;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 34 AGVAVATLV 42
 ID 705 AGVAVATLV 713
 DB
 RESULT 27
 AAE19181
 ID AAE19181 standard, Protein, 918 AA.
 XX
 AC AAE19181:
 XX
 DT 21-MAY-2002 (first entry)
 DE Human protease, PRTS-18 protein.
 XX
 KW Human: protease; PRTS-18; enzyme; gastritis; cirrhosis; Crohn's disease;
 KW gastrointestinal disorder; autoimmune; inflammatory; cell proliferative;
 KW cardiovascular; developmental; epithelial; neurological; reproductive;
 KW AIDS; Acquired Immune Deficiency Syndrome; allergy; rheumatoid arthritis;
 KW anaemia; asthma; atherosclerosis; hypertension; myocardial infarction;
 KW hepatitis; cancer; psoriasis; Cushing's syndrome; hypothyroidism; eczema;
 KW epilepsy; Alzheimer's disease; Huntington's disease; Parkinson's disease;
 KW Pick's disease; infertility; vitiligo; drug screening; gene therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..22
 FT /label= Signal_peptide
 FT Protein 23..918
 FT /label= Mature_PRTS_18_protein
 FT Domain 700..721
 FT /label= Transmembrane_domain
 XX
 FN W200208196 A2.
 PD 31-JAN-2002.
 XX
 XX 17-JUL-2001: 2001WC-0522397.
 PF
 XX 21-JUL-2000: 2000US-220063P
 PR

PR 28-JUL-2000: 2000US-221680P.
 PR 04 AUG 2000: 2000US 223544P.
 PR 11-AUG-2000: 2000US-224717P.
 PR 14-AUG-2000: 2000US-225988P.
 PR 24-AUG-2000: 2000US-227568P.
 XX
 PA (INCYTE GENOMICS INT).
 XX
 PI DeLencastre AM, Gandhi AR, Hatfield AA, Lu DM, Patterson C,
 PI Trilivley CM, Das D, Kalluck DA, Nguyen DB, Lee SA, Khan FA,
 PI Yue H, Au-Yang J, Griffin JA, Policky JL, Rammakar J, Yang J;
 PI Shengavala K, Ding L, Kearney L, Baughn MR, Borowsky ML;
 PI Sanjanwalla MS, Yao MG, Burford N, Wallia NK, Lal Y, Lee S, Todd S;
 PI Lo JP, Tang YL, Elliott VS, Azimzal Y, Lu Y;
 XX
 WP: 2002-479761/51.
 DR N-PSDB: AAB9150.
 XX
 CC New human protease polypeptide, useful in diagnosis, prevention and
 CC treatment of gastrointestinal, cardiovascular, autoimmune/inflammatory,
 CC cell proliferative, developmental, epithelial and neurological
 CC disorders
 CC
 PS Claim 1: Page 158-160; 182pp; English.
 XX
 CC The invention relates to an isolated human protease polypeptide (PRTS).
 CC PRTS protein and DNA are useful for diagnosing, treating and preventing
 CC gastrointestinal disorders (gastritis, cirrhosis, Crohn's disease),
 CC autoimmune/inflammatory disorders (AIDS, allergy, rheumatoid arthritis,
 CC anaemia, asthma), cardiovascular disorder (atherosclerosis, hypertension,
 CC myocardial infarction), cell proliferative disorders (hepatitis, cancer,
 CC psoriasis), developmental disorders (Cushing's syndrome, hypothyroidism),
 CC epithelial disorder (vitiligo, keloid, eczema), neurological disorders
 CC (epilepsy, Alzheimer's disease, Pick's disease, Huntington's disease,
 CC Parkinson's disease), and reproductive disorders (infertility). PRTS
 CC protein is useful in a number of drug screening techniques and to
 CC analyse the proteome of a tissue or cell type. PRTS DNA is useful for
 CC creating knock-in humanised animals or transgenic animals to model human
 CC diseases. In somatic or germ-line gene therapy and in microarrays
 CC utilising fluids or tissues from patients to detect altered PKIN
 CC expression. The present sequence is human PRTS-18 protein.
 CC
 SQ Sequence 918 AA:
 Query Match: 2.8%, Score 9, EB 23, Length 918;
 Best Local Similarity 100.0%, Pred. No. 11;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 34 AGVAVATLV 42
 ID 705 AGVAVATLV 713
 DB
 RESULT 28
 AAE24350
 ID AAE24350 standard; Protein; 926 AA.
 XX
 AC AAE24350:
 XX
 DT 04 OCT 2002 (first entry)
 DE Human novel protease #1.
 XX
 KW Human: novel human protein, NHP; protease; meltrin-beta; ADAM 19;
 KW metalloproteinase; drug screening; clinical trial monitoring; cosmetic;
 KW pharmacological, nutritional, gene therapy, chromosome 5.
 XX
 OS Homo sapiens.
 XX
 FN W0200246759 A2.
 PD 10-MAY-2002.
 XX

BE 40-OCT-2001; 2001WO-0548498.
 XX
 PR 01-NOV-2000; 2000US-244939P.
 PR (LEXI-) LEXICON GENETICS INC.
 PA
 XX
 P1 Walke DW, Wilganowski NL, Friddle CJ;
 P2
 P3 WPI: 2002-479761/51
 P4 N-PSDB; AAD39159.
 P5
 P6
 P7 New nucleic acid encoding a protein that shares sequence similarity
 P8 with mammalian meltrin-beta\ADAM 19 homology metalloprotease for use in
 P9 diagnosis and treatment, and in drug screening .
 P0
 PS Claim 1; Page 35-37; 48pp; English.
 CC The invention relates to novel human proteins (NHP), novel human
 CC proteases that shares sequence similarity with mammalian meltrin-beta\
 CC ADAM 19 homologue metalloprotease, and their corresponding nucleic acid
 CC sequences. NHP DNA is useful for diagnosis, drug screening, clinical
 CC trial monitoring, treatment of diseases and disorders and
 CC pharmacological, cosmetic and nutraceutical applications. It is also
 CC useful as hybridisation probes for screening libraries, and assessing
 CC gene expression patterns (particularly using a microarray or high-
 CC throughput chip format). It is also useful in restriction fragment
 CC length polymorphism (RFLP) analysis to identify specific individuals,
 CC and as a probe to identify the corresponding mutant allele in the
 CC libraries. It is also useful for the detection of mutant
 CC NHPs or inappropriately expressed NHPs for the diagnosis of disease.
 CC It is also useful in gene therapy. The present amino acid sequence
 CC is novel human protease. Human NHP gene is located at chromosome 5.
 CC
 SQ Sequence 926 AA:
 QY 34 AGVLAIVLV 42
 DB 705 AGVLAIVLV 713
 Query Match 2.88; Score 9; DB 23; Length 926;
 Best Local Similarity 100.0%; Prod. No. 11;
 Matches 9; Conservative 0; Mismatches 0. Indels 0. Gaps 0.

DR N-PSDB; AAD39162.
 XX
 PT New nucleic acid encoding a protein that shares sequence
 PT with mammalian meltrin-beta\ADAM 19 homology metallopro
 PT tease for use in diagnosis and treatment, and in drug screening .
 P0
 PS Claim 1; Page 44-47; 48pp; English.
 CC The invention relates to novel human proteins (NHP), novel
 CC proteases that shares sequence similarity with mammal
 CC ADAM 19 homologue metalloprotease, and their correspond
 CC sequences. NHP DNA is useful for diagnosis, drug screen
 CC trial monitoring, treatment of diseases and disorders
 CC pharmacological, cosmetic and nutraceutical applications
 CC useful as hybridisation probes for screening libraries
 CC gene expression patterns (particularly using a microarra
 CC y or high-throughput chip format). It is also useful in restriction
 CC fragment length polymorphism (RFLP) analysis to identify specific
 CC individuals, and as a probe to identify the corresponding mutant allele
 CC in the libraries. It is also useful for the detection of mutant
 CC NHPs or inappropriately expressed NHPs for the diagnosis
 CC of disease. It is also useful in gene therapy. The present amino acid
 CC sequence is novel human protease. Human NHP gene is located at c
 CC
 SQ Sequence 955 AA:
 QY 34 AGVLAIVLV 42
 DB 705 AGVLAIVLV 713
 Query Match 2.88; Score 9; DB 23; Length 9
 Best Local Similarity 100.0%; Prod. No. 11;
 Matches 9; Conservative 0; Mismatches 0. Indels 0. Gaps 0.

RESULT 30
 AAE24352
 ID AAE24352 standard; Protein: 964 AA.
 AC AAE24352;
 DT 04-OCT-2002 (first entry)
 DE Human novel protease #3.
 DE
 XX Human; novel human protein; NHP; protease; meltrin-beta
 KM metalloprotease; drug screening; clinical trial monitor
 KM pharmacological; nutraceutical; gene therapy; chromosome
 OS Homo sapiens.
 XX
 PN WO200146759-A2.
 PD 10-MAY-2002.
 XX
 PF 30-OCT-2001; 2001WO-0548498.
 PF 01-NOV-2000; 2000US-244939P.
 PF (LEXI-) LEXICON GENETICS INC.
 PA
 XX Walke DW, Wilganowski NL, Friddle CJ;
 P1
 P2 WPI: 2002-479761/51.
 P3 N-PSDB; AAD39161.
 P4
 P5 New nucleic acid encoding a protein that shares sequence
 PT with mammalian meltrin-beta\ADAM 19 homology metallopro
 PT tease for use in diagnosis and treatment, and in drug screening .
 P0
 PS Claim 1; Page 41-43; 48pp; English.
 CC The invention relates to novel human proteins (NHP), to

proteases that shares sequence similarity with mammalian serine beta-
ADAM 19 homologue metalloproteinase, and their corresponding amino acid
sequences. NHP DNA is useful for diagnosis, drug screening, clinical
trial monitoring, treatment of diseases and disorders and
pharmacological, cosmetic and nutritional applications. It is also
used as hybridisation probes for screening libraries, and assessing
gene expression patterns (particularly using a microarray or high-
throughput chip format). It is also useful in restriction fragment
length polymorphism (RFLP) analysis to identify specific individuals,
and as a probe to identify the corresponding mutant allele in the
libraries. It is also useful for the detection of mutant
NHPs or inappropriately expressed NHPs for the diagnosis of disease.
It is also useful in gene therapy. The present amino acid sequence
is novel human protease human NHP gene is located at chromosome 5
XX
Sequence 963 AA:
Query Match 2.8%; Score 9; DB 23; Length 963;
Best Local Similarity 100.0%; Pred. No. 11,
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
YY 34 AGVAVATIV 42
DB 705 AGVAVATIV 713
RESULT 31
AAV72965
ID AAV72965 standard; peptide: 36 AA.
AC AAV72965;
XX
DT 13-JUN-2001 (first entry)
XX
DE Mouse type II integral membrane protein, CIRE-like fragment #17.
XX
KW Mouse; type II integral membrane protein; CIRE; anti-allergic; vaccine;
KW immunosuppressive; autoimmune disorder; allergy; viral infection; APC;
KW antigen presenting cell; bacterial infection; macrophage; dendritic cell;
KW DC.
XX
OS Mus musculus.
XX
PN W0200119869-A1.
XX
PD 22-MAR-2001.
XX
PE 13-SEP-2000; 2000MO-AU01096.
XX
PF 13-SEP-1999; 99AU-0002789.
XX
PK (COORD-) COINTEL QUEENSLAND INST MEDICAL RES.
XX
PA Caminschi L, Wright MD, Shortman KD;
XX
PI WPI: 2001-244788/25.
XX
DR N-PSDB; AAD03244.
XX
PT Novel type II integral membrane protein, CIRE, preferentially expressed
PT in dendritic cells, macrophages and their precursors stimulatory to T
PT cells; useful for identifying compounds having immunoregulatory
PT activity.
XX
PS Disclosure: Page 27; 53pp; English.
XX
CC The present sequence is mouse CIRE-like fragment derived from a protein
CC expressed on dendritic cells which has similar function to CIRE protein.
CC The invention relates to mouse type II integral membrane protein, CIRE,
CC and its corresponding cDNA molecule which is preferentially expressed in
CC dendritic cells, macrophages and their precursors stimulatory to T cells.
CC CIRE is useful for screening immunoregulatory compounds and as a marker
CC for dendritic cell (DC) population. CIRE specific ligand such as
CC monoclonal antibodies are useful for isolating an antigen presenting cell

(Ab-2). From a biological sample, CIRE DNA is also used as tools to analyse
the properties and functions of the CIRE gene/protein. For example, the
CIRE DNA molecules are used to generate animal models, which lack functional
CIRE genes and to isolate regulatory regions of the CIRE gene. Such
regulatory regions are used to selectively express exogenous genes in DC
or Ab-2. The invention also relates to methods for modulating an immune
response in a subject by binding and inhibiting the function of an APC
such as a myeloid DC. The CIRE-specific ligands are useful for modulating
immune response by interfering with the function, migration or maturation
of DC or APC. The ligands bind to and interfere with the function of
myeloid dendritic cells such that antigen processing is undertaken by
lymphoid DC which leads to immune suppression and thereby for treating
allergies and autoimmune disorders. The ligands are also useful to block
or enhance the interaction of viruses or bacteria with DC, for improving
resistance to infection. These ligands are also used to target molecules
such as vaccine components, to DC or APC.
XX
Sequence 36 AA:
Query Match 2.5%; Score 8; DB 23; Length 36;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
YY 253 WMGISDLN 260
DB 15 WMGISDLN 22
RESULT 32
AAV72966
ID AAV72966 standard; peptide: 36 AA.
AC AAV72966;
XX
DT 13-JUN-2001 (first entry)
XX
DE Mouse type II integral membrane protein, CIRE-like fragment #18.
XX
KW Mouse; type II integral membrane protein; CIRE; anti-allergic; vaccine;
KW immunosuppressive; autoimmune disorder; allergy; viral infection; APC;
KW antigen presenting cell; bacterial infection; macrophage; dendritic cell;
KW DC.
XX
OS Mus musculus.
XX
PN W0200119869-A1.
XX
PD 22-MAR-2001.
XX
PE 13-SEP-2000; 2000MO-AU01096.
XX
PF 13-SEP-1999; 99AU-0002789.
XX
PK (COORD-) COINTEL QUEENSLAND INST MEDICAL RES.
XX
PA Caminschi L, Wright MD, Shortman KD;
XX
PI WPI: 2001-244788/25.
XX
DR N-PSDB; AAD03244.
XX
PT Novel type II integral membrane protein, CIRE, preferentially expressed
PT in dendritic cells, macrophages and their precursors stimulatory to T
PT cells; useful for identifying compounds having immunoregulatory
PT activity.
XX
PS Disclosure: Page 27; 53pp; English.
XX
CC The present sequence is mouse CIRE-like fragment derived from a protein
CC expressed on dendritic cells which has similar function to CIRE protein.
CC The invention relates to mouse type II integral membrane protein, CIRE,
CC and its corresponding cDNA molecule which is preferentially expressed in
CC dendritic cells, macrophages and their precursors stimulatory to T cells.
CC CIRE is useful for screening immunoregulatory compounds and as a marker

for dendritic cell (DC) population. CIRE specific ligand such as
monoclonal antibodies are useful for isolating an antigen presenting cell
(APC) from a biological sample. CIRE DNA is also used as locus to analyse
the properties and functions of the CIRE gene/protein. For example, the
CIRE genes are used to generate animal models, which lack functional
CIRE genes and to isolate regulatory regions of the CIRE gene. Such
regulatory regions are used to selectively express exogenous genes in DC
or APC. The invention also relates to methods for modulating an immune
response in a subject by binding and inhibiting the function of an APC
such as a myeloid DC. The CIRE-specific ligands are useful for modulating
immune response by interfering with the function, migration or maturation
of DC or APC. The ligands bind to and interfere with the function of
myeloid dendritic cells such that antigen processing is undertaken by
lymphoid DC which leads to immune suppression and anergy for treating
allergies and autoimmune disorders. The ligands are also useful to block
or enhance the interaction of viruses or bacteria with DC, for improving
resistance to infection. These ligands are also used to target molecules
such as vaccine components, to DC or APC.

dendritic cells, macrophages and their precursors stimulate
CIRE is useful for screening immunoregulatory compounds
for dendritic cell (DC) population. CIRE specific ligand
monoclonal antibodies are useful for isolating an antigen
(APC) from a biological sample. CIRE DNA is also used as
the properties and functions of the CIRE gene/protein.
DNA molecules are used to generate animal models, which
CIRE genes and to isolate regulatory regions of the CIRE
regulatory regions are used to selectively express exogenous
or APC. The invention also relates to methods for modulating
response in a subject by binding and inhibiting the function
such as a myeloid DC. The CIRE-specific ligands are useful
immune response by interfering with the function, migration
of DC or APC. The ligands bind to and interfere with the
myeloid dendritic cells such that antigen processing is
lymphoid DC which leads to immune suppression and anergy
allergies and autoimmune disorders. The ligands are also
or enhance the interaction of viruses or bacteria with DC
resistance to infection. These ligands are also used to
such as vaccine components, to DC or APC.

Query Match 2.5%; Score 8; DB 22; Length 36.
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 WVDSPPLS 274
| | | | | | | |
Db 29 WVDSPPLS 36

RESULT 34
AAV72963
ID AAV72963 standard; peptide: 111 AA.

AAV72963;

13-JUN-2001 (first entry)

Mouse type II integral membrane protein, CIRE-like fragment #15.

Mouse, type II integral membrane protein, CIRE, antiallergic, vaccine;
immunosuppressive, autoimmune disorder, allergy; viral infection; APC;
antigen presenting cell, bacterial infection, macrophage, dendritic cell,
DC.

Mus musculus.

W0200119869-A1.

22-MAR-2001.

13-SEP-2000; 2000WO-AU01096.

13-SEP-1999; 99AU-0002788.

(COOUN-) COOUNCIL QUEBENS LAND INST MEDICAL RES.

Caminschi I, Wright MD, Shortman KD;

WPI: 2001-244788/25.

N-PSOH; AAD03242.

Novel type II integral membrane protein, CIRE, preferentially expressed
in dendritic cells, macrophages and their precursors stimulatory to T
cells, useful for identifying compounds having immunoregulatory
activity.

Disclosure: Page 26; 53pp; English.

The present sequence is mouse CIRE-like fragment derived from a protein
expressed on dendritic cells which has similar function to CIRE protein
CC The invention relates to mouse type II integral membrane protein, CIRE
and its corresponding cDNA molecule which is preferentially expressed in

Sequence 111 AA;

Query Match 2.5%; Score 8; DB 22; Length 1
Best Local Similarity 100.0%; Pred. No. 4;
Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY 292 EDCAFFSG 299
| | | | | | | |
Db 77 EDCAFFSG 84

RESULT 34
AAG26292
ID AAG26292 standard; Protein: 212 AA.

AAG26292;

17-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 30694.
Protein identification: signal transduction pathway; ac
hybridisation assay; genetic mapping; gene expression co
termination sequence.

Arabidopsis thaliana.

EP1033405-A2.

06-SEP-2000.

25-FEB-2000; 2000PP-0301449.

25-FEB-1999; 9905-0121845.

05-MAR-1999; 9905-0123180.

09-MAR-1999; 9905-0123548.

23-MAR-1999; 9905-0125784.

25-MAR-1999; 9905-0126784.

29-MAR-1999; 9905-0126784.

01-APR-1999; 9905-0127462.

06-APR-1999; 9905-0128244.

08-APR-1999; 9905-0128714.

16-APR-1999; 9905-0129845.

19-APR-1999; 9905-0130077.

21-APR-1999; 9905-0130449.

23-APR-1999; 9905-0130510.

28-APR-1999; 9905-0130891.

30-APR-1999; 9905-0131449.

40-APR-1999; 9905-0132407.

04-MAY-1999; 9905-0132484.

05-MAY-1999; 9905-0132485.

PR 06-MAY-1999; 9905-0132486;
PR 06-MAY-1999; 9905-0132487;
PR 07-MAY-1999; 9905-0132863;
PR 11-MAY-1999; 9905-0134256;
PR 14-MAY-1999; 9905-0134218;
PR 14-MAY-1999; 9905-0134219;
PR 14-MAY-1999; 9905-0134221;
PR 14-MAY-1999; 9905-0134370;
PR 18-MAY-1999; 9905-0134768;
PR 19-MAY-1999; 9905-0134941;
PR 20-MAY-1999; 9905-0135121;
PR 21-MAY-1999; 9905-0135353;
PR 24-MAY-1999; 9905-0135629;
PR 25-MAY-1999; 9905-0136021;
PR 27-MAY-1999; 9905-0136392;
PR 28-MAY-1999; 9905-0136782;
PR 01-JUN-1999; 9905-0137222;
PR 03-JUN-1999; 9905-0137528;
PR 04-JUN-1999; 9905-0137502;
PR 07-JUN-1999; 9905-0137724;
PR 08-JUN-1999; 9905-0138094;
PR 10-JUN-1999; 9905-0138540;
PR 10-JUN-1999; 9905-0138847;
PR 14-JUN-1999; 9905-0139119;
PR 16-JUN-1999; 9905-0139452;
PR 16-JUN-1999; 9905-0139453;
PR 17-JUN-1999; 9905-0139492;
PR 18-JUN-1999; 9905-0139454;
PR 18-JUN-1999; 9905-0139455;
PR 18-JUN-1999; 9905-0139456;
PR 18-JUN-1999; 9905-0139457;
PR 18-JUN-1999; 9905-0139458;
PR 18-JUN-1999; 9905-0139459;
PR 18-JUN-1999; 9905-0139460;
PR 18-JUN-1999; 9905-0139461;
PR 18-JUN-1999; 9905-0139462;
PR 18-JUN-1999; 9905-0139463;
PR 18-JUN-1999; 9905-0139750;
PR 18-JUN-1999; 9905-0139763;
PR 21-JUN-1999; 9905-0139817;
PR 22-JUN-1999; 9905-0139899;
PR 23-JUN-1999; 9905-0140454;
PR 23-JUN-1999; 9905-0140354;
PR 24-JUN-1999; 9905-0140695;
PR 28-JUN-1999; 9905-0140823;
PR 29-JUN-1999; 9905-0140991;
PR 30-JUN-1999; 9905-0141287;
PR 01-JUL-1999; 9905-0141842;
PR 02-JUL-1999; 9905-0142055;
PR 06-JUL-1999; 9905-0142390;
PR 08-JUL-1999; 9905-0142803;
PR 09-JUL-1999; 9905-0142920;
PR 12-JUL-1999; 9905-0143427;
PR 13-JUL-1999; 9905-0143542;
PR 14-JUL-1999; 9905-0143624;
PR 15-JUL-1999; 9905-0144005;
PR 16-JUL-1999; 9905-0144085;
PR 16-JUL-1999; 9905-0144086;
PR 19-JUL-1999; 9905-0144325;
PR 19-JUL-1999; 9905-0144331;
PR 19-JUL-1999; 9905-0144332;
PR 19-JUL-1999; 9905-0144333;
PR 19-JUL-1999; 9905-0144334;
PR 19-JUL-1999; 9905-0144335;
PR 20-JUL-1999; 9905-0144352;
PR 20-JUL-1999; 9905-0144632;
PR 23-JUL-1999; 9905-0144884;
PR 23-JUL-1999; 9905-0144814;
PR 21-JUL-1999; 9905-0145086;
PR 21-JUL-1999; 9905-0145088;
PR 22-JUL-1999; 9905-0145085;
PR 22-JUL-1999; 9905-0145087;

PR 22-JUL-1999; 9905-0145089;
PR 22-JUL-1999; 9905-0145192;
PR 23-JUL-1999; 9905-0145145;
PR 23-JUL-1999; 9905-0145218;
PR 23-JUL-1999; 9905-0145224;
PR 26-JUL-1999; 9905-0145276;
PR 27-JUL-1999; 9905-0145913;
PR 27-JUL-1999; 9905-0145918;
PR 27-JUL-1999; 9905-0145919;
PR 28-JUL-1999; 9905-0145951;
PR 02-AUG-1999; 9905-0146386;
PR 02-AUG-1999; 9905-0146388;
PR 02-AUG-1999; 9905-0146389;
PR 03-AUG-1999; 9905-0147038;
PR 04-AUG-1999; 9905-0147204;
PR 04-AUG-1999; 9905-0147302;
PR 05-AUG-1999; 9905-0147192;
PR 05-AUG-1999; 9905-0147260;
PR 06-AUG-1999; 9905-0147303;
PR 06-AUG-1999; 9905-0147416;
PR 09-AUG-1999; 9905-0147493;
PR 09-AUG-1999; 9905-0147945;
PR 10-AUG-1999; 9905-0148171;
PR 11-AUG-1999; 9905-0148319;
PR 12-AUG-1999; 9905-0148341;
PR 13-AUG-1999; 9905-0148565;
PR 13-AUG-1999; 9905-0148684;
PR 16-AUG-1999; 9905-0149368;
PR 17-AUG-1999; 9905-0149175;
PR 18-AUG-1999; 9905-0149426;
PR 20-AUG-1999; 9905-0149723;
PR 20-AUG-1999; 9905-0149929;
PR 23-AUG-1999; 9905-0149902;
PR 23-AUG-1999; 9905-0149949;
PR 25-AUG-1999; 9905-0150566;
PR 26-AUG-1999; 9905-0150884;
PR 27-AUG-1999; 9905-0151065;
PR 27-AUG-1999; 9905-0151066;
PR 27-AUG-1999; 9905-0151080;
PR 30-AUG-1999; 9905-0151303;
PR 31-AUG-1999; 9905-0151438;
PR 01-SEP-1999; 9905-0151940;
PR 07-SEP-1999; 9905-0152363;
PR 10-SEP-1999; 9905-0153070;
PR 13-SEP-1999; 9905-0153758;
PR 15-SEP-1999; 9905-0154018;
PR 16-SEP-1999; 9905-0154039;
PR 20-SEP-1999; 9905-0154779;
PR 22-SEP-1999; 9905-0155139;
PR 23-SEP-1999; 9905-0155486;
PR 24-SEP-1999; 9905-0155659;
PR 28-SEP-1999; 9905-0156458;
PR 29-SEP-1999; 9905-0156596;
PR 04-OCT-1999; 9905-0157117;
PR 05-OCT-1999; 9905-0157753;
PR 06-OCT-1999; 9905-0157865;
PR 07-OCT-1999; 9905-0158029;
PR 08-OCT-1999; 9905-0158232;
PR 12-OCT-1999; 9905-0158369;
PR 13-OCT-1999; 9905-0159293;
PR 13-OCT-1999; 9905-0159294;
PR 13-OCT-1999; 9905-0159295;
PR 14-OCT-1999; 9905-0159329;
PR 14-OCT-1999; 9905-0159330;
PR 14-OCT-1999; 9905-0159331;
PR 14-OCT-1999; 9905-0159637;
PR 14-OCT-1999; 9905-0159638;
PR 18-OCT-1999; 9905-0159584;
PR 21-OCT-1999; 9905-0160741;
PR 21-OCT-1999; 9905-0160767;
PR 21-OCT-1999; 9905-0160768;
PR 21-OCT-1999; 9905-0160770;

Query Match	2.5%	Score R	FR 21	Length 212
Best Local Similarity	100.0%	Pred. No. 25		
Matches 8	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY 26.2 EGTQWWD 269				
Db 15 EGTQWWD 22				
RESUL1 15				
AA626291				
ID AA626291 standard, Protein; 214 AA.				
XX AA626291:				
XX				
DI 17-OCT-2000 (first entry)				
XX				
DE Arabidopsis thaliana protein fragment SFG ID No: 30693				
XX				
KW Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.				
KW				
XX Arabidopsis thaliana.				
XX				
PN EP1033405-A2.				
XX				
PD 06-SEP-2000.				
XX				
XX 25-FEB-2000; 2000EP-0301439.				
XX				
PR 25-FEB-1999; 990S-0121825.				
PR 05-MAR-1999; 990S-0123180.				
PR 09-MAR-1999; 990S-0123548.				
PR 23-MAR-1999; 990S-0125788.				
PR 25-MAR-1999; 990S-0126564.				
PR 29-MAR-1999; 990S-0126785.				
PR 01-APR-1999; 990S-0127462.				
PR 06-APR-1999; 990S-0128234.				
PR 08-APR-1999; 990S-0128714.				
PR 16-APR-1999; 990S-0129845.				
PR 19-APR-1999; 990S-0130077.				
PR 21-APR-1999; 990S-0130449.				
PR 23-APR-1999; 990S-0130510.				
PR 23-APR-1999; 990S-0130891.				
PR 28-APR-1999; 990S-0131449.				
PR 30-APR-1999; 990S-0132048.				
PR 30-APR-1999; 990S-0132407.				
PR 04-MAY-1999; 990S-0132484.				
PR 05-MAY-1999; 990S-0132485.				
PR 06-MAY-1999; 990S-0132486.				
PR 06-MAY-1999; 990S-0132487.				
PR 07-MAY-1999; 990S-0142864.				
PR 11-MAY-1999; 990S-0134256.				
PR 14-MAY-1999; 990S-0134218.				
PR 14-MAY-1999; 990S-0134219.				
PR 14-MAY-1999; 990S-0134221.				

PR	14-MAY-1999	99005-0134776
PR	14-MAY-1999	99005-0134778
PR	13-MAY-1999	99005-0134941
PR	20-MAY-1999	99005-0135124
PR	21-MAY-1999	99005-0135363
PR	24-MAY-1999	99005-0135629
PR	25-MAY-1999	99005-0136021
PR	27-MAY-1999	99005-0136372
PR	28-MAY-1999	99005-0136788
PR	01-JUN-1999	99005-0137262
PR	03-JUN-1999	99005-0137588
PR	04-JUN-1999	99005-0137828
PR	07-JUN-1999	99005-0137744
PR	08-JUN-1999	99005-0138094
PR	10-JUN-1999	99005-0138510
PR	10-JUN-1999	99005-0138847
PR	14-JUN-1999	99005-0139191
PR	16-JUN-1999	99005-0139462
PR	16-JUN-1999	99005-0139463
PR	17-JUN-1999	99005-0139476
PR	18-JUN-1999	99005-0139474
PR	18-JUN-1999	99005-0139478
PR	18-JUN-1999	99005-0139481
PR	18-JUN-1999	99005-0139482
PR	18-JUN-1999	99005-0139483
PR	18-JUN-1999	99005-0139484
PR	18-JUN-1999	99005-0139485
PR	18-JUN-1999	99005-0139486
PR	18-JUN-1999	99005-0139487
PR	18-JUN-1999	99005-0139488
PR	18-JUN-1999	99005-0139489
PR	18-JUN-1999	99005-0139490
PR	18-JUN-1999	99005-0139491
PR	18-JUN-1999	99005-0139492
PR	18-JUN-1999	99005-0139493
PR	18-JUN-1999	99005-0139494
PR	18-JUN-1999	99005-0139495
PR	18-JUN-1999	99005-0139496
PR	18-JUN-1999	99005-0139497
PR	18-JUN-1999	99005-0139498
PR	18-JUN-1999	99005-0139499
PR	18-JUN-1999	99005-0139500
PR	18-JUN-1999	99005-0139501
PR	18-JUN-1999	99005-0139502
PR	18-JUN-1999	99005-0139503
PR	18-JUN-1999	99005-0139504
PR	18-JUN-1999	99005-0139505
PR	18-JUN-1999	99005-0139506
PR	18-JUN-1999	99005-0139507
PR	18-JUN-1999	99005-0139508
PR	18-JUN-1999	99005-0139509
PR	18-JUN-1999	99005-0139510
PR	18-JUN-1999	99005-0139511
PR	18-JUN-1999	99005-0139512
PR	18-JUN-1999	99005-0139513
PR	18-JUN-1999	99005-0139514
PR	18-JUN-1999	99005-0139515
PR	18-JUN-1999	99005-0139516
PR	18-JUN-1999	99005-0139517
PR	18-JUN-1999	99005-0139518
PR	18-JUN-1999	99005-0139519
PR	18-JUN-1999	99005-0139520
PR	18-JUN-1999	99005-0139521
PR	18-JUN-1999	99005-0139522
PR	18-JUN-1999	99005-0139523
PR	18-JUN-1999	99005-0139524
PR	18-JUN-1999	99005-0139525
PR	18-JUN-1999	99005-0139526
PR	18-JUN-1999	99005-0139527
PR	18-JUN-1999	99005-0139528
PR	18-JUN-1999	99005-0139529
PR	18-JUN-1999	99005-0139530
PR	18-JUN-1999	99005-0139531
PR	18-JUN-1999	99005-0139532
PR	18-JUN-1999	99005-0139533
PR	18-JUN-1999	99005-0139534
PR	18-JUN-1999	99005-0139535
PR	18-JUN-1999	99005-0139536
PR	18-JUN-1999	99005-0139537
PR	18-JUN-1999	99005-0139538
PR	18-JUN-1999	99005-0139539
PR	18-JUN-1999	99005-0139540
PR	18-JUN-1999	99005-0139541
PR	18-JUN-1999	99005-0139542
PR	18-JUN-1999	99005-0139543
PR	18-JUN-1999	99005-0139544
PR	18-JUN-1999	99005-0139545
PR	18-JUN-1999	99005-0139546
PR	18-JUN-1999	99005-0139547
PR	18-JUN-1999	99005-0139548
PR	18-JUN-1999	99005-0139549
PR	18-JUN-1999	99005-0139550
PR	18-JUN-1999	99005-0139551
PR	18-JUN-1999	99005-0139552
PR	18-JUN-1999	99005-0139553
PR		


```

PR 27-AUG-1999; 990S-0151065-
PR 27-AUG-1999; 990S-0151066-
PR 30-AUG-1999; 990S-0151080-
PR 30-AUG-1999; 990S-0151303-
PR 31-AUG-1999; 990S-0151438-
PR 01-SEP-1999; 990S-0151940-
PR 07-SEP-1999; 990S-0152363-
PR 10-SEP-1999; 990S-0153070-
PR 13-SEP-1999; 990S-0153788-
PR 15-SEP-1999; 990S-0154018-
PR 16-SEP-1999; 990S-0154039-
PR 20-SEP-1999; 990S-0154779-
PR 22-SEP-1999; 990S-0155139-
PR 23-SEP-1999; 990S-0155486-
PR 24-SEP-1999; 990S-0155553-
PR 28-SEP-1999; 990S-0156458-
PR 29-SEP-1999; 990S-0156596-
PR 04-OCT-1999; 990S-0157117-
PR 05-OCT-1999; 990S-0157753-
PR 06-OCT-1999; 990S-0157865-
PR 07-OCT-1999; 990S-0158029-
PR 08-OCT-1999; 990S-0158332-
PR 12-OCT-1999; 990S-0158369-
PR 13-OCT-1999; 990S-0159293-
PR 13-OCT-1999; 990S-0159294-
PR 13-OCT-1999; 990S-0159295-
PR 14-OCT-1999; 990S-0159329-
PR 14-OCT-1999; 990S-0159331-
PR 14-OCT-1999; 990S-0159637-
PR 14-OCT-1999; 990S-0159638-
PR 18-OCT-1999; 990S-0159584-
PR 21-OCT-1999; 990S-0160741-
PR 21-OCT-1999; 990S-0160767-
PR 21-OCT-1999; 990S-0160768-
PR 21-OCT-1999; 990S-0160770-
PR 21-OCT-1999; 990S-0160814-
PR 21-OCT-1999; 990S-0160815-
PR 22-OCT-1999; 990S-0160846-
PR 22-OCT-1999; 990S-0160981-
PR 22-OCT-1999; 990S-0160989-
PR 25-OCT-1999; 990S-0161404-
PR 25-OCT-1999; 990S-0161405-
PR 25-OCT-1999; 990S-0161406-
PR 25-OCT-1999; 990S-0161454-
PR 26-OCT-1999; 990S-0161360-
PR 26-OCT-1999; 990S-0161361-
PR 28-OCT-1999; 990S-0161920-
PR 28-OCT-1999; 990S-0161992-
PR 28-OCT-1999; 990S-0161993-
PR 29-OCT-1999; 990S-0162142

```

```

Query Match: 2.5%, Score 8, DB 21, Length 279;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 262 ECTMOWVD 269
DB 82 ECTMOWVD 89

RESULT 38
AA088024
ID AA088024 standard, Protein: 325 AA.
XX
AC AA088024;
XX
DI 05-JUN-2002 (first entry)
XX
DE Mouse protein encoded by Con14 1A.
XX
KW Nucleic acid library; immune response; asthma; COPD;
airway hyperresponsiveness; bronchiolovular manifestation;

```

```

KW signature sequence; SS; chronic obstructive pulmonary disease;
KW allergic disease; rhinitis; atopic dermatitis; urticaria;
KW autoimmune disease; multiple sclerosis; inflammatory bowel disease;
KW allograft rejection; infectious disease.
KW calcium-activated chloride channel.
XX
XX Mus sp.
XX W020214366.A2.
XX
XX 21-FEB-2002.
XX
XX 16 AUG 2001; 2001W0-N100610.
XX
XX 16 AUG 2000; 2000P-0202867.
XX
XX (UTUT-) RIKSUNIV UTRECHT.
XX
XX Groot PC, Van Bergenhenegouwen BJ, Van Oosterhout ADM;
XX WPL, 2000; 21982790
XX N FSDA, ABK47192.
XX
XX Nucleic acid library comprising genes which are capable of initiation,
XX proliferation and suppression of an immune response, especially an immune
XX response observed with airway hyper-responsiveness or asthma
XX
XX Example 10, Fig 10; 120pp; English.
XX
XX The invention relates to a nucleic acid library comprising genes or
XX their fragments which are capable of modulating an immune response
XX observed with airway hyperresponsiveness and/or bronchiolovular
XX manifestations of asthma. Also included are a method for modulating an
XX immune response of an individual comprising modulating a gene comprising
XX a nucleic acid at least functionally equivalent to a nucleic acid
XX identifiable by a signature sequence (SS) given in the specification such
XX as RI-SO-R1 A11, St01-A10, S302-1 C11, St01-A12, and RI-SO-R1-B7, a
XX substance (for use as a medicament) capable of modulating a gene
XX comprising a nucleic acid at least functionally equivalent to a nucleic
XX acid identifiable by SS and the use of a proteaceous substance derived
XX from a nucleic acid at least functionally equivalent to a nucleic acid
XX identifiable by SS for the production of an antagonist (for use as a
XX medicament) against the substance. The antagonist and substance are
XX useful for the treatment of an immune response observed with airway
XX hyper-responsiveness and/or bronchiolovular manifestations of asthma.
XX The method is useful for modulating the above immune response, where the
XX gene encodes a gene product capable of modulating the immune response.
XX The substance is useful for treating an immune response, particularly
XX asthma, chronic obstructive pulmonary disease (COPD), allergic diseases
XX (rhinitis, atopic dermatitis, urticaria), autoimmune diseases (e.g.
XX multiple sclerosis), inflammatory bowel disease, allograft rejection and
XX infectious disease. The present sequence is a mouse or human
XX protein encoded by a signature sequence gene or its homologue/functionally
XX equivalent.
XX
XX Sequence 125 AA;

```

```

Query Match: 2.5%, Score 8, DB 23, Length 325;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 23 IVMLQLSP 30
DB 54 IVMLQLSP 61

RESULT 39
AA088026
ID AA088026 standard, Protein: 125 AA.
XX
AC AA088026;
XX
DI 05-JUN 2002 (first entry)
XX
KW

```

XX Human DC-SIGN protein.
 DE
 XX
 XX Nucleic acid library: immune response; asthma; COPD;
 KM airway hyperresponsiveness; bronchoalveolar manifestation;
 KM signature sequence; SS: chronic obstructive pulmonary disease;
 KM allergic disease; rhinitis; atopic dermatitis; urticaria;
 KM allograft rejection; infectious disease; inflammatory bowel disease;
 KM calcium-activated chloride channel
 XX
 OS Homo sapiens.
 XX
 PN W0200214366-A2.
 XX
 PD 21-FEB-2002.
 XX
 PF 16-AUG-2001; 2001WO-NL00610.
 XX
 PK 16-AUG-2001; 2001EP-0203467
 XX
 PA (UWI-) PIJCSUNIV UTPRECHT.
 XX
 PI Groof PC, Van Bergenhenegouwen BJ, Van Oosterhout AJM;
 DR WPI: 2002-241888/29.
 XX
 PT Nucleic acid library comprising genes which are capable of initiation,
 PT progression and suppression of an immune response, especially an immune
 PT response observed with airway hyper-responsiveness of asthma
 XX
 PS Example 10: Fig 11: 120pp: English.
 XX
 CC The invention relates to a nucleic acid library comprising genes or
 CC their fragments which are capable of modulating an immune response
 CC observed with airway hyperresponsiveness and/or bronchoalveolar
 CC manifestations of asthma. Also included are a method for modulating an
 CC immune response of an individual comprising modulating a gene comprising
 CC a nucleic acid at least functionally equivalent to a nucleic acid
 CC identifiable by a signature sequence (SS) given in the specification such
 CC as RI-SO-RI-A11, StOI-A10, SVO2-1-C11, StOI-A12, and RI-SO-RI-B7, a
 CC substance (for use as a medicament) capable of modulating a gene
 CC comprising a nucleic acid at least functionally equivalent to a nucleic
 CC acid identifiable by SS and the use of a proteolaceous substance derived
 CC from a nucleic acid at least functionally equivalent to a nucleic acid
 CC identifiable by SS for the production of an antagonist (for use as a
 CC medicament) against the substance. The antagonist and substance are
 CC useful for the treatment of an immune response observed with asthma.
 CC hyperresponsiveness and/or bronchoalveolar manifestations of asthma.
 CC The method is useful for modulating the above immune response, where the
 CC gene encodes a gene product capable of modulating the immune response.
 CC The substance is useful for treating an immune response, particularly
 CC asthma, chronic obstructive pulmonary disease (COPD), allergic diseases
 CC (rhinitis, atopic dermatitis, urticaria), autoimmune diseases (e.g.
 CC multiple sclerosis), inflammatory bowel disease, allograft rejection and
 CC infectious disease. The present sequence is a mouse or human
 CC protein encoded by a signature sequence gene or its homologous/functional
 CC equivalent.
 XX
 XX Sequence 125 AA:
 QY
 DB 23 LVQLLSF 30
 1111111111
 54 LVQLLSF 61
 RESULT 40
 ABP25448
 ID ABP25448 standard; Protein: 356 AA.

XX
 AC ABP25448;
 XX
 XX 02-JUL-2002 (first entry)
 DT
 XX Streptococcus polypeptide SEQ ID NO 72.
 XX
 DE Streptococcus; GAS; GBS, group B streptococcus; Streptococcus
 KM group A streptococcus; Streptococcus pyogenes; antibiotic
 KM antiinflammatory; infection; vaccine; meningitis; gene
 XX
 XX Streptococcus pyogenes.
 OS
 XX W0200234771-A2.
 XX
 PN 02-MAY-2002.
 XX
 PD 29-OCT-2001; 2001WO-GH04789.
 XX
 PF 27-OCT-2001; 2003GB-0026344.
 XX
 PK 24-NOV-2000; 2000GB-0028747.
 XX
 PR 07-MAR-2001; 2001GB-0005640.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Telford J, Masiganti V, Margalit KOS YI, Grandi G,
 PI Telford H;
 DR WPI: 2002-352546/38.
 XX
 DR N-PSDR: ARN66074.
 XX
 PT New Streptococcus protein for the treatment or prevention of
 PT or disease caused by Streptococcus bacteria, such as new
 PT for detecting a compound that binds to the protein -
 XX
 PS Claim 1: Page 3161; 4525pp: English.
 XX
 CC The invention relates to a protein (ABP25413-ABP30895) of
 CC Streptococcus/GBS (Streptococcus agalactiae) or group A
 CC (Streptococcus pyogenes), comprising one of 5483 sequences
 CC in the specification. The proteins have antibacterial and
 CC activity. (1), nucleic acids encoding (1), ARN66044-ARN71;
 CC antibodies that bind (1) are used in the manufacture of
 CC the treatment or prevention of infection or disease caused
 CC Streptococcus bacteria, particularly S. agalactiae and S.
 CC Nucleic acids encoding (1) are used to detect Streptococcus
 CC biological sample. (1) is used to determine whether a
 CC (1). A composition comprising (1) or a nucleic acid
 CC used as a vaccine or diagnostic composition. The disease
 CC Streptococcus that is prevented or treated may be meningitis,
 CC acid encoding (1) may be used to recombinantly produce (1)
 CC used in gene therapy. Antibodies to (1) are used for
 CC chromatography, immunoassays, and disinfectant/identifying
 CC Streptococcus proteins.
 XX
 XX Sequence 356 AA:
 QY
 DB 178 QELTDLKT 185
 1111111111
 78 QELTDLKT 85
 RESULT 41
 AA081511
 ID AA081511 standard; Protein: 472 AA.
 XX
 AC AA081511;
 XX

09-APR-2002 (first entry)
 C. elegans interflagellar transport protein, IFT52.
 Human, chlamydomonas, intraflagellar transport protein, Cht-2.
 anti-bacterial; protozoacide; contraceptive; anti-inflammatory; malaria;
 Chlamydomonas, IFT, mouse, kidney disease; retinal disorder; nematode;
 thyroid disorder; chondrocyte disease; olfactory disease; anosmia;
 primary ciliary dyskinesia, insect, protozoa, male contraceptive,
 parasitic disease; African sleeping sickness; trypanosomiasis;
 trichomoniasis; giardiasis; flagella.
 Caenorhabditis elegans.
 W0200190307-AZ.
 29-NOV-2001.
 24-MAY-2001; 2001WO-0517104.
 24-MAY-2000; 200003-2069234.
 (UYMA-) UNIV MASSACHUSETTS.
 Wilman GB, Pacour CJ, Rosenbaum JL, Cole DC,
 WPI: 2002 089926/12.
 Novel Chlamydomonas or mouse intraflagellar transport (IFT) protein or
 its variant, useful for identifying modulators that are useful for
 treating nematode, insect, protozoa or bacterial infection by
 inhibiting IFT.
 Disclosures: Figure 9D; 132PF; English.
 The invention relates to an isolated polypeptide (:) of *C. elegans*
 intraflagellar transport (IFT) particle protein, Cht-2, or mouse IFT
 particle protein 57. Also described are methods which are useful for:
 diagnosing kidney disease, retinal disorder, thyroid disorder,
 chondrocyte disease, olfactory disease, anosmia, or primary ciliary/
 dyskinesia; for treating an infection in a mammal (a human) or plant
 caused by nematode, insect, protozoa or bacteria. The nucleic acid
 molecules are useful for diagnosis of disorders associated with aberrant
 expression of nucleic acid molecules and in genetic mapping and
 chromosome identification. The polypeptides are useful in generation of
 antibodies, as reagents in diagnostic assays, for the identification of
 other cellular gene products or compounds that can modulate the activity
 or expression of nucleic acids or polypeptides, and as pharmaceutical
 reagents useful for the treatment of disorders associated with aberrant
 expression or activity of the nucleic acids or polypeptides. IFT can
 serve as a male contraceptive. The IFT proteins can be targeted for
 treating parasitic diseases such as malaria, African sleeping sickness,
 trypanosomiasis, leishmaniasis, trichomoniasis, and giardiasis. An anti-IFT
 drug, taken orally, would inhibit assembly of flagella in newly divided
 Giardia and cause disassembly of previously formed flagella in non-
 dividing Giardia. An anti-IFT drug would block assembly of the flagellum
 and the flagellar sheath and affect the trypanosome's life cycle. The
 anti-IFT modulators can also be used to treat trichomonos vaginitis in
 humans. IFT inhibitors are also useful for combating phytopathogenic
 nematodes inhibiting insect pest by blocking sensory cilia function
 and/or assembly, leaving an insect unable to smell, unable to taste and
 unable to hear. AA081494-AA081529 represent intraflagellar transport
 particle amino acid sequences of the invention.
 Sequence 472 AA:
 Query Match. 2.5%, Score 8, DB 23, Length 472;
 Best Local Similarity 100.0%, Freq. No. 32;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 169 DQSK0001 176
 1111111111
 21 DQSK0001 28

RESULT 42
 AA095815
 ID AA095815 standard; Protein: 596 AA.
 AA095815:
 25-JUN-2001 (first entry)
 Human protein sequence SMO ID NO:18813.
 Human, primer, detection, diagnosis; antisense therapy; gene therapy.
 Homo sapiens.
 EP1074617-AZ.
 07-FEB-2001.
 28-JUN-2000; 2000EP-0116126.
 29-JUN-1999; 99JP-0248036.
 27-AUG-1999; 99JP-0300253.
 11-7AN-2000; 2000JP-0118776.
 02-MAY-2000; 2000JP-0183767.
 09-JUN-2000; 2000JP-0241899.
 (HELI-) HELIX RES INST.
 Chta T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 Ishii S, Shojiyama T, Wakamatsu A, Nagai K, Otsuki T;
 WPI: 2001-318749/34.
 Primer sets for synthesizing polynucleotides, particularly the 5602
 full-length cDNAs defined in the specification, and for the detection
 and/or diagnosis of the abnormality of the proteins encoded by the
 full-length cDNAs.
 Claim 8, SMO ID 18813, 2537PF, CD RCM; English.
 The present invention describes primer sets for synthesizing 5602
 full-length cDNAs defined in the specification. Where a primer set
 comprises (a) an oligo-dT primer and an oligonucleotide complementary
 to the complementary strand of a polynucleotide which comprises one of
 the 5602 nucleotide sequences defined in the specification, where the
 oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 of an oligonucleotide comprising a sequence complementary to the
 complementary strand of a polynucleotide which comprises a 5'-end
 sequence and an oligonucleotide comprising a sequence complementary to a
 polynucleotide which comprises a 3'-end sequence, where the
 oligonucleotide comprises at least 15 nucleotides and the combination of
 the 5'-end sequence/3'-end sequence is selected from those defined in
 the specification. The primer sets can be used in antisense therapy and
 in gene therapy. The primers are useful for synthesizing polynucleotides,
 particularly full-length cDNAs. The primers are also useful for the
 detection and/or diagnosis of the abnormality of the proteins encoded by
 the full-length cDNAs. The primers allow obtaining of the full-length
 cDNAs easily without any specialised methods. AA003146 to AA013628 and
 AA013633 to AA018742 represent human cDNA sequences; AA092446 to
 AA098893 represent human amino acid sequences; and AA013629 to AA013632
 represent oligonucleotides, all of which are used in the exemplification
 of the present invention.
 Sequence 596 AA:
 Query Match. 2.5%, Score 8, DB 22, Length 596;
 Best Local Similarity 100.0%, Freq. No. 65;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 43 QVSKVPS 50
 1111111111

Db 558 QVSKVPSS 565

RESULT 43

AAV24094 standard; Protein: 763 AA.

AAV24094.

04-SEP-1999 (first entry)

Mouse Epc1.

Epc1, Epc2, enhancer of polycomb, tumour suppressor, diagnosis; proliferative disorder; developmental disorder; neoplasia.

Mus sp.

W0943986-A1.

08-JUL-1999.

28-DEC-1998; 98W0-0527667.

10-DEC-1997; 97US-0068957.

(CHIR) CHIRON CORP.

Randazzo F.

WPI: 1999-419108/35.

N-PSDB: AAX83993.

Mammalian tumour suppressor, termed enhancer of polycomb, useful for therapy, prognosis and diagnosis of proliferative and developmental disorders

Claim 2; Page 49; 55pp; English.

The present sequence represents a mouse tumour suppressor, designated enhancer of polycomb (Epc).

Human Epc genes can be used to identify a

10-11 region of human chromosome 10 and a q13.3 region of human

chromosome 22. The Epc genes and expression products are useful as

therapeutic, prognostic and diagnostic tools for proliferative and

developmental disorders, e.g. neoplasia.

Query Match 2.58, Score 8, Db 20, Length 763.

Best Local Similarity 100.0%; Pred. No. 81;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 43 QVSKVPSS 50

Db 725 QVSKVPSS 732

RESULT 44

AAH45133

AAH45133 standard; Protein: 27 AA.

AAH45133.

12-FEB-2001 (first entry)

Human secreted protein sequence encoded by gene 14 Spg in No-74

Human; secreted protein; immunosuppressive; antiarthritic; arthritic; antiproliferative; cytostatic; cardiac; vasotropic; cerebroprotective; neuroprotective; antibacterial; virucide; fungicide; cancer; opthalmological; autoimmune disease; hyperproliferative disorder; cardiovascular disorder; cerebrovascular disorder; wound healing; nervous system disorder; aging; chemotaxis.

XX Homo sapiens.

XX W0200058467-A1.

XX 05-OCT-2000.

XX 24-MAR-2000; 2000W0-0507505.

XX 26-MAR-1999; 99US-0126502.

XX 17-DEC-1999; 99US-0172410.

XX (HUMA-) HUMAN GENOME SCI INT.

XX Rosen CA, Ruben SM, Komatsu S.

XX WPI: 2000-611712/58.

XX N-PSDB: AAC80544.

XX Nucleic acid molecules encoding human secreted proteins

preventing, treating or ameliorating a disorder, e.g. Al

Parkinson's diseases and cancers

Claim 11; Page 385; 440pp; English.

Polynucleotide sequences AAC80544-C80580 represent cDNA

secreted proteins AAB45120-145169. Sequences AAB45170-145

alternative polypeptides encoded by the genes, and amino

to which they are homologous. The genes and proteins have

dependent on the tissues and cells in which they are exp

of their activities include immunosuppressive; antiarth

antiinflammatory; antiproliferative; cytostatic; cardiac;

cerebroprotective; neuroprotective; antibacte

fungicide; and ophthalmological. The secreted proteins, p

antagonists and agonists may be useful in treating, prev

diagnosing diseases and disorders such as autoimmune dis

e.g. rheumatoid arthritis, hyperproliferative disorders

of the breast or liver, cardiovascular disorders e.g. car

cerebrovascular disorders e.g. cerebral ischemia, and/or

system disorders e.g. Alzheimer's disease, infections ca

bacteria, viruses and fungi and ocular disorders e.g. cor

The polypeptides can also be used to aid wound healing a

cell proliferation, to prevent skin aging due to sunburr

organs before transplantation, for supporting cell cultu

tissues, to regenerate tissues and in chemotaxis. The p

also be used as a food additive or preservative to incre

storage capabilities. AAC80522-C80530 and AAB45119 repr

used in the isolation and characterisation of the genes

the invention.

Query Match 2.49, Score 7, Db 21, Length 2

Best Local Similarity 100.0%; Pred. No. 34;

Matches 7; Conservative 0; Mismatches 0; Indels 0;

UY 32 LAGVIV 38

Db 20 LAGVIV 26

RESULT 45

AAR22106

AAR22106 standard; Protein: 27 AA.

AAR22106.

15-JUL-1992 (first entry)

Peptide 3-29 derived from CD18 leukocyte integrin.

Cell adhesion; intercellular adhesion molecule; endothel

inflammation; leukocyte chemotaxis; rhinovirus; common

```

XX      W9204473-A.
PN
XX
XX      05-MAR-1992.
FD
XX      24-AUG-1991;      91WO-0006063.
PF
XX      27-AUG-1990;      90US-0573624.
PR
XX      (CETU ) CETUS CORP.
PA
XX      Liu DY, Kaymakçalan Z, Mundy K:
PI
XX      WPI; 1992-096833/12.
DR
XX      Peptide(s) derived from beta sub-unit CD18 of leukocyte integrins
PT      - prevent leukocyte binding to ICAM and leukocyte chemotaxis, for
PR      treating inflammatory diseases and rhinoviral infection
XX
XX      Claim 2; Page 23; 31pp; English.
PS
XX
XX      This peptide (derived from CD18 see AAR24256) was synthesised and
CC      tested for capacity to interfere with or block adhesion of
CC      polymorphonuclear leukocytes to human endothelial cell monolayers.
CC      The peptide shows inhibitory activity at 10(power -4)M. The peptide
CC      can be used to treat/prevent inflammation and can inhibit binding
CC      of viruses to endothelium. See AAQ22780 and AAR22104 R22112.
XX
XX      Sequence      29 AA:
SU
XX
XX      Query Match      2.2%; Score 7; DB 13; Length 29;
XX      Best Local Similarity 100.0%; Pred. No. 35;
XX      Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      71 ANGEISE 77
XX      1111111
DB      5 ANGEISE 11

```

Search completed: December 7, 2002, 11:15:26
 Job time : 40 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd

OM protein - protein search, using sw model

Run on: December 7, 2002, 11:14:40 : Search time 17 seconds
(without alignments)
562,497 Million cell updates/sec

Title: US-09-831-458A-12

Perfect score: 325
Sequence: 1 MDSKEPRVQUGLIGLACIGH...RCVNDNWICKKPAACEFIDE 325

Scoring table:
Gapop 60.0, Gapext 60.0

Searchbed: 262574 seqs, 26422922 residues

Word size: 5

Total number of hits satisfying chosen parameters: 6756

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: Issued Patents AA:

- 1: /cgn2_6/prodata/1/1aa/5A.COMB.pep:*
- 2: /cgn2_6/prodata/1/1aa/5B.COMB.pep:*
- 3: /cgn2_6/prodata/1/1aa/6A.COMB.pep:*
- 4: /cgn2_6/prodata/1/1aa/6B.COMB.pep:*
- 5: /cgn2_6/prodata/1/1aa/PCITUS.COMB.pep:*
- 6: /cgn2_6/prodata/1/1aa/Backfillal.pep:*

Pred. No is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	168	51.7	296	4	US-09-517-605-17 Sequence 17, Appl
2	78	24.0	494	4	US-09-517-605-2 Sequence 2, Appl
3	9	2.8	751	2	US-08-836-443-3 Sequence 3, Appl
4	7	2.2	154	1	US-08-150-945-67 Sequence 67, Appl
5	7	2.2	154	4	US-08-976-161-67 Sequence 67, Appl
6	7	2.2	352	5	PCT-US96-01314-58 Sequence 58, Appl
7	7	2.2	353	5	PCT-US96-01314-53 Sequence 53, Appl
8	7	2.2	548	2	US-09-632-312-4 Sequence 4, Appl
9	7	2.2	548	2	US-08-993-318A-4 Sequence 4, Appl
10	7	2.2	548	4	US-09-999-886-4 Sequence 4, Appl
11	7	2.2	548	4	US-09-396-260-4 Sequence 4, Appl
12	7	2.2	548	4	US-09-576-281-4 Sequence 10, Appl
13	7	2.2	676	3	US-08-630-172-10 Sequence 10, Appl
14	7	2.2	676	4	US-09-375-419-10 Sequence 10, Appl
15	7	2.2	720	3	US-09-257-799-48 Sequence 48, Appl
16	7	2.2	720	3	US-08-920-919A-48 Sequence 48, Appl
17	7	2.2	769	4	US-08-789-078-1 Sequence 1, Appl
18	7	2.2	769	2	US-08-752-633-1 Sequence 1, Appl
19	7	2.2	769	2	US-08-476-062A-45 Sequence 45, Appl
20	7	2.2	769	4	US-07-728-115-31 Sequence 31, Appl
21	7	2.2	769	4	US-08-938-085A-31 Sequence 31, Appl
22	7	2.2	769	5	PCT-US95-04886-1 Sequence 45, Appl
23	7	2.2	769	5	PCT-US95-01314-45 Sequence 45, Appl
24	6	1.8	8	6	US-08-480-236C-12 Sequence 12, Appl
25	6	1.8	9	6	US-08-480-236C-12 Sequence 12, Appl
26	6	1.8	14	2	US-08-659-235C-12 Sequence 12, Appl
27	6	1.8	14	2	US-08-659-235C-12 Sequence 12, Appl

28	6	1.8	19	1	US-07-908-417-10 Appl
29	6	1.8	19	5	PCT-US94-06171-19 Appl
30	6	1.5	25	4	US-08-819-624-24 Appl
31	6	1.8	25	4	US-09-150-812-24 Appl
32	6	1.8	33	4	US-08-849-624-23 Appl
33	6	1.8	33	4	US-09-150-812-23 Appl
34	6	1.8	37	4	US-08-849-624-25 Appl
35	6	1.8	37	4	US-09-150-812-25 Appl
36	6	1.8	39	1	US-08-189-831-143 Appl
37	6	1.8	39	2	US-08-471-066-143 Appl
38	6	1.8	40	4	US-08-689-421-1 Appl
39	6	1.8	40	4	US-09-489-528-1 Appl
40	6	1.8	40	4	US-09-181-827A-1 Appl
41	6	1.8	49	4	US-08-849-624-26 Appl
42	6	1.8	49	4	US-09-150-812-26 Appl
43	6	1.8	52	4	US-08-946-165A-808 Appl
44	6	1.8	58	1	US-07-664-989B-78 Appl
45	6	1.8	58	1	US-08-458-160-55 Appl

ALIGNMENTS

RESULT 1

US-09-517-605-17
Sequence 17, Application US/09/517605
Patent No. 6391567
GENERAL INFORMATION:
APPLICANT: Liltman, Dan R.
APPLICANT: Kwon, Douglas S.
APPLICANT: van Kooyk, Yvette
APPLICANT: Geljtenbeck, Ineo
TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RE-
FILE REFERENCE: 1049-1-017
CURRENT AFFILIATION NUMBER: US/09/517,605
CURRENT FILING DATE: 2000-03-02
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 17
LENGTH: 296
TYPE: PRT
ORGANISM: Homo sapiens
US-09-517-605-17

Query Match 51.7% Score 168; DB 4; Length:
Best Local Similarity 100.0%; Pred. No. 7.5e-154;
Matches 168, Conservative 0, Mismatches 0, Indels 0

QY 1 MDSKEPRVQUGLIGLACIGHALVQISPMIAGVVALVSVSPSS
DB 1 MDSKEPRVQUGLIGLACIGHALVQISPMIAGVVALVSVSPSS
QY 61 IYQNLIDLKAAGVPEISKSLQETVQELTQKAAGVPEISKSLQETVQ
DB 61 IYQNLIDLKAAGVPEISKSLQETVQELTQKAAGVPEISKSLQETVQ
QY 121 IPRRSKIQFYQPTPIKAAVGPPIPKSKIQETVQETPIKAAVGPPI
DB 121 IPRRSKIQFYQPTPIKAAVGPPIPKSKIQETVQETPIKAAVGPPI

RESULT 2
US-09-517-605-2
Sequence 2, Application US/09/517605
Patent No. 6391567
GENERAL INFORMATION:
APPLICANT: Liltman, Dan R.
APPLICANT: Kwon, Douglas S.
APPLICANT: van Kooyk, Yvette
APPLICANT: Geljtenbeck, Ineo
TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RE-

FILE REFERENCE: 1049-1-017
 CURRENT APPLICATION NUMBER: US/09/517,605
 CURRENT FILING DATE: 2000-03-02
 NUMBER OF SEQ ID NOS: 17
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO: 2
 LENGTH: 404
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-517-605-2

Query Match 24.0%; Score 78; DB 4; Length 404;
 Best Local Similarity 100.0%; Pred. No. 4,1e-67;
 Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QDAIVQNTQLKAAGVLSKSKTLEIYQELTQLKAAGVLPKSKLOEIVQELTPKAA 117
 DB 74 QDAIVQNTQLKAAGVLSKSKTLEIYQELTQLKAAGVLPKSKLOEIVQELTPKAA 117
 QY 118 VGEIPEKSKLOEIVQELT 135
 DB 134 VGEIPEKSKLOEIVQELT 151

RESULT 3
 US-08-836-443-3
 Sequence 3; Application US/08836443
 Patent No. 5883241
 GENERAL INFORMATION:
 APPLICANT: DOHERTY, Andrew, J.P.
 APPLICANT: SLOCUMBE, Patrick, M.
 TITLE OF INVENTION: DNA SEQUENCES CODING FOR A
 TITLE OF INVENTION: HUMAN METALLOPROTEINASE AND VARIANTS THEREOF
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESS: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
 STREET: 130 Water Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/836,443
 FILING DATE: 01-MAY-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 9704420
 FILING DATE: 05-SEP-1997
 APPLICATION NUMBER: 9612145, 4
 FILING DATE: 11-JUN-1996
 APPLICATION NUMBER: 9526230, 9
 FILING DATE: 21-DEC-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Resnick, David S.
 REGISTRATION NUMBER: 34,235
 REFERENCE/DOCKET NUMBER: 47424
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-523-1400
 TELEFAX: 617-523-6440
 TELEX:
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 751 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: NO. 5883241e
 US-08-836-443-3

Query Match 2.8%; Score 7; DB 2; Length 751;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 14 ASYLVAILV 42
 DB 540 ASYLVAILV 538

RESULT 4
 US-08-450-945-67
 Sequence 67; Application US/08450945
 Patent No. 5763383
 GENERAL INFORMATION:
 APPLICANT: Kondo, Kazuhito
 APPLICANT: Mocarski, Edward S., Jr.
 TITLE OF INVENTION: LATENT TRANSPLANTS AND PROMOTERS
 TITLE OF INVENTION: OF CYTOMEGALOVIRUS
 NUMBER OF SEQUENCES: 75
 CORRESPONDENCE ADDRESS:
 ADDRESS: Dehlinger & Associates
 STREET: 350 Cambridge Avenue, Suite 250
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94306
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/450,945
 FILING DATE: 23-MAY-1995
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: Sholtz, Charles K.
 REGISTRATION NUMBER: 38,615
 REFERENCE/DOCKET NUMBER: 8600-0157
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 324-0880
 TELEFAX: (415) 324-0960
 INFORMATION FOR SEQ ID NO: 67:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 154 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-450-945-67

Query Match 2.8%; Score 7; DB 1; Length 154;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 GVLVAILV 41
 DB 23 GVLVAILV 29
 RESULT 5
 US-08-976-161-67
 Sequence 67; Application US/08976161
 Patent No. 6194542
 GENERAL INFORMATION:
 APPLICANT: Kondo, Kazuhito
 APPLICANT: Mocarski, Edward S., Jr.
 TITLE OF INVENTION: LATENT TRANSPLANTS AND PROMOTERS
 TITLE OF INVENTION: OF CYTOMEGALOVIRUS
 NUMBER OF SEQUENCES: 75
 CORRESPONDENCE ADDRESS:
 ADDRESS: Dehlinger & Associates
 STREET: 350 Cambridge Avenue, Suite 250

CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-08/976,161
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/450,945
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 8600-0157
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 154 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-976-161-67

Query Match 2.28 Score 71 DB 4 Length 154
Best Local Similarity 100.0% Pred. No. 35;
Matches 71 Conservative 0 Mismatches 0 Indels 0 Gaps 0

QY 35 GVLVALI.41
|||||

DB 23 GVLVALI.29

RESULT 6
PCT-US96-01314-58
Sequence 58, Application PC/TUS9601314
GENERAL INFORMATION:
APPLICANT: M. Amin Arnaout
TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
NUMBER OF SEQUENCES: 78
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: Wordperfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/01314
FILING DATE: 30-JAN-96
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/380,167
FILING DATE: 30-JAN-95
ATTORNEY/AGENT INFORMATION:
NAME: John W. Freeman
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/267001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906

TELEX: 200154
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-01314-58

Query Match 2.28 Score 71 DB 5 Length 252
Best Local Similarity 100.0% Pred. No. 56;
Matches 71 Conservative 0 Mismatches 0 Indels 0 Gaps 0

QY 71 AVGEISE 77
|||||

DB 232 AVGEISE 238

RESULT 7
PCT-US96-01314-53
Sequence 53, Application PC/TUS9601314
GENERAL INFORMATION:
APPLICANT: M. Amin Arnaout
TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
NUMBER OF SEQUENCES: 78
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: Wordperfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/01314
FILING DATE: 30-JAN-96
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/480,167
FILING DATE: 30-JAN-95
ATTORNEY/AGENT INFORMATION:
NAME: John W. Freeman
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/267001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-01314-53

Query Match 2.28 Score 71 DB 5 Length 253
Best Local Similarity 100.0% Pred. No. 56;
Matches 71 Conservative 0 Mismatches 0 Indels 0 Gaps 0

QY 71 AVGEISE 77
|||||

DB 233 AVGEISE 239

RESULT 8
US-09-032-315-4

Sequence 4, Application US/09032315
Patent No. 5985818
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Xu, Feng
TITLE OF INVENTION: LACCASE MUTANTS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5985818, No. 5985818, No. 5985818, No. 5985818, No. 5985818, No. 5985818, No. 5985818, No. 5985818, No. 5985818, No. 5985818
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/032,315
FILING DATE: 27-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Rozek, Carol
REGISTRATION NUMBER: 36,993
REFERENCE/DOCKET NUMBER: 5200,200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-032-315-4
Query Match 2.2% Score 7, DB 2, Length 548,
Best Local Similarity 100.0%, Freq. No. 1.26+02,
Matches 7, Conservative 0, Mismatches 0, Indels 0, Gaps 0,
UY 308 DWDNYWI 314
Db 275 DWDNYWI 281
RESULT 9
US-08-993-318A-4
Sequence 4, Application US/08993318A
Patent No. 5998353
GENERAL INFORMATION:
APPLICANT: Pedersen, Anders
APPLICANT: Svendsen, Allan
APPLICANT: Schneider, Palle
APPLICANT: Rasmussen, Gertel
APPLICANT: Cherry, Joel
TITLE OF INVENTION: LACCASE MUTANTS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5998353, No. 5998353, No. 5998353, No. 5998353, No. 5998353, No. 5998353, No. 5998353, No. 5998353, No. 5998353, No. 5998353
STREET: 405 Lexington Avenue
CITY: New York
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,318A

FILING DATE: December 18, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gregg, Valeria A.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 5032,200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-993-318A-4
Query Match 2.2% Score 7, DB 2, Length 548,
Best Local Similarity 100.0%, Freq. No. 1.26+02,
Matches 7, Conservative 0, Mismatches 0, Indels 0, Gaps 0,
UY 308 DWDNYWI 314
Db 275 DWDNYWI 281
RESULT 10
US-09-399-886-4
Sequence 4, Application US/09399886
Patent No. 6140092
GENERAL INFORMATION:
APPLICANT: Pedersen, Anders
APPLICANT: Svendsen, Allan
APPLICANT: Schneider, Palle
APPLICANT: Rasmussen, Gertel
APPLICANT: Cherry, Joel
TITLE OF INVENTION: LACCASE MUTANTS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6140092, No. 6140092, No. 6140092, No. 6140092, No. 6140092, No. 6140092, No. 6140092, No. 6140092, No. 6140092, No. 6140092
STREET: 405 Lexington Avenue
CITY: New York
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/399,886
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,318
FILING DATE: December 18, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Gregg, Valeria A.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 5032,200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-09-349-886-4

Query Match
Best Local Similarity 100.0%; Score 7, DB 4, Length 548,
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 DVDNYWI 314
DB 275 DVDNYWI 281

RESULT 11

US-09-349-260-4
Sequence 4, Application US/09396260
Patent No. 6184015
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Xu, Feng
TITLE OF INVENTION: LACCASE MUTANTS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6184015, No. 6184015, 1st of No. 6184015, America, Inc.,
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/396,260
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/032,315
FILING DATE: 27-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Rozek, Carol
REGISTRATION NUMBER: 36,993
REFERENCE/DOCKET NUMBER: 5200, 200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-09-349-260-4
Query Match
Best Local Similarity 100.0%; Score 7, DB 4, Length 548,
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 DVDNYWI 314
DB 275 DVDNYWI 281

RESULT 12

US-09-576-281-4
Sequence 4, Application US/09576281
Patent No. 6277611
GENERAL INFORMATION:
APPLICANT: Peperisen, Anders
APPLICANT: Svendsen, Allan
APPLICANT: Schneider, Palle
APPLICANT: Kasmussen, Grethe

APPLICANT: Cherry, Joel
TITLE OF INVENTION: LACCASE MUTANTS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6277611, No. 6277611, 1st of No. 6277611,
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (Free)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/576,281
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,418
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Gregg, Valeria A.
REGISTRATION NUMBER: 34,728
REFERENCE/DOCKET NUMBER: 6042, 200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0124
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-09-576-281-4
Query Match
Best Local Similarity 100.0%; Score 7, DB 4, Length 548,
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 DVDNYWI 314
DB 275 DVDNYWI 281

RESULT 13

US-08-630-172-10
Sequence 10, Application US/08630172
Patent No. 6060054
GENERAL INFORMATION:
APPLICANT: Staerz, Dwe
TITLE OF INVENTION: NOVEL PRODUCT AND PROCESS FOR IT
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & M. Thirush
STREET: 1700 Lincoln Street, 45th Floor
CITY: Denver
STATE: Colorado
COUNTRY: U.S.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,172
FILING DATE:
CLASSIFICATION: 51A
ATTORNEY/AGENT INFORMATION:

NAME: Connell, Gary J.
PRIOR APPLICATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2879-36
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 676 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-630-172-10

Query Match: 2.28; Score 7; DB 3; Length 676;
Best Local Similarity 100.0%; Prod No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 71 AVGELSE 77
|||||
DB 319 AVGELSE 325

RESULT 14
US-09-375-419-10
Sequence 10, Application US/09375419
Patent No. 6264950
GENERAL INFORMATION:
APPLICANT: Staerz, Uwe
TITLE OF INVENTION: NOVEL PRODUCT AND PROCESS FOR T
TITLE OF INVENTION: LYMPHOCYTE VETO
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, 35th Floor
CITY: Denver
STATE: Colorado
COUNTRY: U.S.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/375,419
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/630 172
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
PRIOR APPLICATION NUMBER: 32,020
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 676 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-475-419-10

Query Match: 2.28; Score 7; DB 4; Length 676;
Best Local Similarity 100.0%; Prod No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 71 AVGELSE 77
|||||
DB 319 AVGELSE 325

RESULT 15
US-09-257-799-48
Sequence 48, Application US/09257799
Patent No. 6020178
GENERAL INFORMATION:
APPLICANT: Aschlimann, Daniel
TITLE OF INVENTION: Transglutaminase and Gene Encoding Same
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/257,799
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/920,919
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Berson, Bennett J.
PRIOR APPLICATION NUMBER: 37094
REFERENCE/DOCKET NUMBER: 960296, 94509
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 720 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-257-799-48

Query Match: 2.28; Score 7; DB 3; Length 720;
Best Local Similarity 100.0%; Prod No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 269 DCSFLSP 275
|||||
DB 551 DCSFLSP 557

RESULT 16
US-08-920-919A-48
Sequence 48, Application US/08920919A
Patent No. 6114119
GENERAL INFORMATION:
APPLICANT: Aschlimann, Daniel
TITLE OF INVENTION: Transglutaminase and Gene Encoding Same
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/920,919A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Betson, Bennett J
REGISTRATION NUMBER: 37094
REFERENCE/DOCKET NUMBER: 960296, 94509
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 720 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-920-919A-48

Query Match 2.28; Score 7; DB 3; Length 720;
Best Local Similarity 100.0%; Pred. No. 1.5e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 269 DCSPLSP 275
DB 551 DCSPLSP 557

RESULT 17
US-08-789-078-1
Sequence 1, Application US/08789078
Patent No. 5843885
GENERAL INFORMATION:
APPLICANT: Benedict, Stephen H
APPLICANT: Shahaan, Teruna
APPLICANT: Chan, Marcia
APPLICANT: Tibbets, Scott
TITLE OF INVENTION: ICAM-1/LFA-1 PEPTIDES FOR INDUCING
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS
STREET: 1101 Walnut St.
CITY: Kansas City
STATE: MO
COUNTRY: USA
ZIP: 64106
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/789,078
FILING DATE: 03-FEB-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/229,513
FILING DATE: 19-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Collins, John M.
REGISTRATION NUMBER: 26262
REFERENCE/DOCKET NUMBER: 22833
TELECOMMUNICATION INFORMATION:
TELEPHONE: (816)474-9050
TELEFAX: 816/474-9057
TELEX: 434-363
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 769 amino acids
TYPE: amino acid

TOPOLOGY: both
MOLECULE TYPE: protein
HYDROTHERICAL: No
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: tonsil
FEATURE:
NAME/KEY: Region
LOCATION: 1..22
OTHER INFORMATION: /label "signal"
OTHER INFORMATION: /note "signal sequence"
FEATURE:
NAME/KEY: Region
LOCATION: 449..496
OTHER INFORMATION: /label "repeat"
OTHER INFORMATION: /note "cysteine rich repeat"
FEATURE:
NAME/KEY: Region
LOCATION: 497..540
OTHER INFORMATION: /label "repeat"
OTHER INFORMATION: /note "cysteine rich repeat"
FEATURE:
NAME/KEY: Region
LOCATION: 541..581
OTHER INFORMATION: /label "repeat"
OTHER INFORMATION: /note "cysteine rich repeat"
FEATURE:
NAME/KEY: Region
LOCATION: 582..617
OTHER INFORMATION: /label "repeat"
OTHER INFORMATION: /note "cysteine rich repeat"
FEATURE:
NAME/KEY: Domain
LOCATION: 701..723
OTHER INFORMATION: /label "transmembrane domain"
OTHER INFORMATION: /note "transmembrane domain"
FEATURE:
NAME/KEY: Domain
LOCATION: 724..769
OTHER INFORMATION: /label "cyto"
OTHER INFORMATION: /note "cytoplasmic domain"
PUBLICATION INFORMATION:
AUTHORS: Pigott,
TITLE: LFA-1 Amino acid sequence (B2) (from human
JOURNAL: The Adhesion Molecule Facts Book
PAGES: 96-96
DATE: 1993
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 769
US-08-789-078-1
Query Match 2.28; Score 7; DB 2; Length 7
Best Local Similarity 100.0%; Pred. No. 1.6e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0;

QY 71 AVGEISL 77
DB 341 AVGEISL 347

RESULT 18
US-08-752-633-1
Sequence 1, Application US/08752633
Patent No. 5853889
GENERAL INFORMATION:
APPLICANT: Benedict, Stephen H.
APPLICANT: Shahaan, Teruna
APPLICANT: Chan, Marcia
APPLICANT: Tibbets, Scott
TITLE OF INVENTION: ICAM-1/LFA-1 PEPTIDES FOR INDUCING
NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:
 ADDRESSER: HOVEY, WILLIAMS, TIMMONS & COLLINS
 STREET: 1101 Walnut St.
 CITY: Kansas City
 STATE: MO
 COUNTRY: USA
 ZIP: 64106
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/752,633
 FILING DATE:
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Collins, John M.
 REGISTRATION NUMBER: 26262
 REFERENCE/DOCKET NUMBER: 22833
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (816)474-9050
 TELEFAX: (816)474-9057
 TELEX: 434-363
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 769 amino acids
 TYPE: amino acid
 TOPOLOGY: both
 MOLECULE TYPE: protein
 HYPOTHEetical: NO
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 TISSUE TYPE: Tonsil
 FEATURE:
 NAME/KEY: Region
 LOCATION: 1..22
 OTHER INFORMATION: /label="signal
 FEATURE:
 NAME/KEY: Region
 LOCATION: 449..496
 OTHER INFORMATION: /label="repeat
 FEATURE:
 NAME/KEY: Region
 LOCATION: 497..540
 OTHER INFORMATION: /label="repeat
 FEATURE:
 NAME/KEY: Region
 LOCATION: 541..581
 OTHER INFORMATION: /label="repeat
 FEATURE:
 NAME/KEY: Region
 LOCATION: 582..617
 OTHER INFORMATION: /label="repeat
 FEATURE:
 NAME/KEY: Domain
 LOCATION: 701..723
 OTHER INFORMATION: /label="trans
 FEATURE:
 NAME/KEY: Domain
 LOCATION: 724..769
 OTHER INFORMATION: /label="cyto
 PUBLICATION INFORMATION:
 AUTHORS: Pigott,
 TITLE: LFA-1 Amino acid sequence (B2) (from human

TITLE: Tonsil
 JOURNAL: The Adhesion Molecule Facts Book
 PAGE: 96-96
 DATE: 1993
 RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 769
 US-08-752-633-1
 Query Match: 2.2% Score 7. DB 2. Length 769;
 Best local similarity 100.0%, Pred. No. 1.6e+02;
 Matches 7: conservative 0: mismatches 0: indels 0:
 CY 71 AVEHSE 77
 DB 341 AVEHSE 447
 RESULT 19
 US-08-476-062A-45
 Sequence 47. Alignment US/02176062A
 Patent No. 5872275
 GENERAL INFORMATION:
 APPLICANT: Attacott, M. Amin
 TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
 NUMBER OF INVENTIONS: RESPONSES WITH BETA2 INTEGRINS
 NUMBER OF SEQUENCES: 53
 CORRESPONDENCE ADDRESS:
 ADDRESSER: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: US
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM compatible
 OPERATING SYSTEM: Windows95
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/98/476,062A
 FILING DATE: 07-JUN-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/216,081
 FILING DATE: 21-MAR-1994
 APPLICATION NUMBER: 07/647,840
 FILING DATE: 04-JAN-1991
 APPLICATION NUMBER: 07/539,842
 FILING DATE: 18-JUN-1990
 APPLICATION NUMBER: 07/212,573
 FILING DATE: 28-JUN-1988
 ATTORNEY/AGENT INFORMATION:
 NAME: Freeman, John W.
 REGISTRATION NUMBER: 29,066
 REFERENCE/DOCKET NUMBER: 06785/068003
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617/542-5070
 TELEFAX: 617/542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 45:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 769 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FRAGMENT TYPE: internal
 US-08-476-062A-45
 Query Match: 2.2% Score 7. DB 2. Length 769;
 Best local similarity 100.0%, Pred. No. 1.6e+02;
 Matches 7: conservative 0: mismatches 0: indels 0:
 CY 71 AVEHSE 77
 DB 341 AVEHSE 447


```

RESULT 20
US-07-728-215-31
: Sequence 31, Application US/07728215
: Patent No. 5962643
: GENERAL INFORMATION:
: APPLICANT: Sheppard, Dean
: APPLICANT: Quaranta, Vito
: TITLE OF INVENTION: A No. 5962643el Integrin Beta Subunit and Uses
: TITLE OF INVENTION: Thereof
: NUMBER OF SEQUENCES: 43
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pretly, Schroeder, Brueggemann & Clark
: STREET: 4370 La Jolla Village Drive, Suite 700
: CITY: San Diego
: STATE: California
: COUNTRY: United States of America
: ZIP: 92122
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/728,215
: FILING DATE: 19910711
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Campbell, Cathryn A.
: REGISTRATION NUMBER: 31,815
: REFERENCE/DOCKET NUMBER: P31 8717
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 535-9001
: TELEFAX: (619) 535-8949
: INFORMATION FOR SEQ ID NO: 31:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 769 amino acids
: TYPE: AMINO ACID
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-07-728-215-31

Query Match 2.2% Score 7: DB 2: Length 769:
Best Local Similarity 100.0%: Pred. No. 1.6e+02:
Matches 7: Conservative 0: Mismatches 0: Incls 0: Gaps 0:

QY 71 AVEELSE 77
DB 341 AVEELSE 347

```

```

RESULT 21
US-08-938-085A-31
: Sequence 31, Application US/08938085A
: Patent No. 6339148
: GENERAL INFORMATION:
: APPLICANT: Sheppard, Dean
: APPLICANT: Quaranta, Vito
: TITLE OF INVENTION: A No. 6339148el Integrin Beta Subunit and Uses
: TITLE OF INVENTION: Thereof
: NUMBER OF SEQUENCES: 62
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Crew LLP
: STREET: Two Embarcadero Center, Eighth Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-3834
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk

```

```

: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: 85704556, 085A
: FILING DATE: 26-SEP-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/728,215
: FILING DATE: 11-JUL-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Parent, Annette S.
: REGISTRATION NUMBER: 42,058
: REFERENCE/DOCKET NUMBER: 024077-080210HS
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 576-0200
: TELEFAX: (415) 576-0300
: INFORMATION FOR SEQ ID NO: 41:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 769 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-938-085A-41

Query Match 2.2% Score 7: DB 4: Length 76
Best Local Similarity 100.0%: Pred. No. 1.6e+02:
Matches 7: Conservative 0: Mismatches 0: Incls 0:

QY 71 AVEELSE 77
DB 341 AVEELSE 347

```

```

RESULT 22
PCT-US95-04886-1
: Sequence 1, Application PC/US9504886
: GENERAL INFORMATION:
: APPLICANT: Benedict, Stephen H.
: APPLICANT: Stahen, Teruna
: APPLICANT: Chan, Marcia
: APPLICANT: Tibbels, Scott
: TITLE OF INVENTION: ICAM-1/LFA-1 PEPTIDES FOR INJECTN
: TITLE OF INVENTION: IMMUNE TOLERANCE
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: HOVEY, WILLIAMS, LIMMONS & COLLINS
: STREET: 1101 Walnut St.
: CITY: Kansas City
: STATE: MO
: COUNTRY: USA
: ZIP: 64106
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/04886
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: COLLINS, John M.
: REGISTRATION NUMBER: 26472
: REFERENCE/DOCKET NUMBER: 22844
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (816)474-9050
: TELEFAX: (816)474-9057
: TELEX: 434-363
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 769 amino acids

```

```

TYPE: amino acid
TOPOLOGY: both
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: tonsil
FEATURE:
NAME/KEY: Region
LOCATION: 1..22
OTHER INFORMATION: /label= signal
OTHER INFORMATION: /note= "signal sequence"
FEATURE:
NAME/KEY: Region
LOCATION: 449..496
OTHER INFORMATION: /label= repeat
OTHER INFORMATION: /note= "cysteine rich repeat"
FEATURE:
NAME/KEY: Region
LOCATION: 497..540
OTHER INFORMATION: /label= repeat
OTHER INFORMATION: /note= "cysteine rich repeat"
FEATURE:
NAME/KEY: Region
LOCATION: 541..581
OTHER INFORMATION: /label= repeat
OTHER INFORMATION: /note= "cysteine rich repeat"
FEATURE:
NAME/KEY: Region
LOCATION: 582..617
OTHER INFORMATION: /label= repeat
OTHER INFORMATION: /note= "cysteine rich repeat"
FEATURE:
NAME/KEY: Domain
LOCATION: 701..723
OTHER INFORMATION: /label= trans
OTHER INFORMATION: /note= "transmembrane domain"
FEATURE:
NAME/KEY: Domain
LOCATION: 724..769
OTHER INFORMATION: /label= cyto
OTHER INFORMATION: /note= "cytoplasmic domain"
PUBLICATION INFORMATION:
AUTHORS: Pigott,
TITLE: LFA-1 Amino acid sequence (B2) (From human
TITLE: tonsil)
JOURNAL: The Adhesion Molecule Facts Book
PAGES: 96-96
DATE: 1993
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 769
PCT-US95-04886-1

Query Match 2.2%; Score 7; DB 5; Length 769;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 71 ANGEISE 77
DB 341 ANGEISE 347

```

```

CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01314
FILING DATE: 30-JAN-96
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/480,167
FILING DATE: 30-JAN-95
ATTORNEY/AGENT INFORMATION:
NAME: John W. Freeman
REGISTRATION NUMBER: 29,066
PREFERENCE/WORK NUMBER: 00784/267001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 769
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
PCT-US96-01314-45

Query Match 2.2%; Score 7; DB 5; Length 769;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 71 ANGEISE 77
DB 341 ANGEISE 347

RESULT 24
5185431-19
Patent No. 5185431
APPLICANT: YOSHIMATSU, KENTARO; SHIKATA, YASUSHI; TANAKA, ISAO;
HASEGAWA, YOSHIKAZU; SETO, TOSHIO; OSAWA, TOSHIO
TITLE OF INVENTION: RECOMBINANT NATURAL KILLER CELL ACTIVATOR
NUMBER OF SEQUENCES: 31
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 05/07/392,841
FILING DATE: 11-AUG-1989
SEQ ID NO: 19
LENGTH: 8
5185431-19

Query Match 1.8%; Score 6; DB 6; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 266 QWYDGS 271
DB 2 QWYDGS 7

RESULT 25
5185431-20
Patent No. 5185431
APPLICANT: YOSHIMATSU, KENTARO; SHIKATA, YASUSHI; TANAKA, ISAO;
HASEGAWA, YOSHIKAZU; SETO, TOSHIO; OSAWA, TOSHIO
TITLE OF INVENTION: RECOMBINANT NATURAL KILLER CELL ACTIVATOR
NUMBER OF SEQUENCES: 31
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 05/07/392,841
FILING DATE: 11-AUG-1989

```

SEQ ID NO:20:
LENGTH: 9
5185431-20

Query Match

1.8%; Score 6; DB 6; Length 9;
Best Local Similarity 100.0%; Pred. No. 26+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 271 SPLSPS 271
|||||

DB 3 SPLSPS 8

RESULT 26
US-08-480-229C-12
Sequence 12, Application US/08480229C
Patent No. 5874562
GENERAL INFORMATION:
APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmunds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0026-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-229C-12

Query Match 1.8%; Score 6; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 271 SPLSPS 276
|||||

DB 1 SPLSPS 6

RESULT 27
US-08-659-235C-12
Sequence 12, Application US/08659235C
Patent No. 5877281
GENERAL INFORMATION:
APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Bridgid

APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmunds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235C
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869 8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-659-235C-12

Query Match 1.8%; Score 6; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 271 SPLSPS 276
|||||

DB 1 SPLSPS 6

RESULT 28
US-07-908-317-10
Sequence 10, Application US/07908317
Patent No. 5420027
GENERAL INFORMATION:
APPLICANT: FISHER, CHARLES W.
APPLICANT: BARNES, HENRY J.
APPLICANT: ESTABROOK, RONALD W.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
TITLE OF INVENTION: THE EXPRESSION OF BIOLOGICALLY
TITLE OF INVENTION: ACTIVE FUSION PROTEINS COMBINE
TITLE OF INVENTION: ERYTHROCYTE PYCNOPLASM 2450 FUSION
TITLE OF INVENTION: A REDUCTASE IN BACTERIA
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & PARKER
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US-07/908,317
FILING DATE: 19920702
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: PARKER, DAVID L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UTSD:292/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512-320-7200
TELEFAX: 512-474-7577
TELEX: NOT APPLICABLE
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acid residues
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
US-07-908-317-10

Query Match 1.8%; Score 6; DB 1; Length 19;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 LSPML 33
|||||
Db 11 LSPML 16

RESULT 29
PCT-US93-06171-10
Sequence 10, Application PCT/US9306171
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: FUSION PROTEINS COMPRISING
TITLE OF INVENTION: ENKARYOTIC CYTOCHROME P450 FUSFD 53
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOID, WHITE & DURREE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06171
FILING DATE: 19930629
CLASSIFICATION:
PRIOR APPLICATION NUMBER:
APPLICATION DATA:
FILING DATE: 07/908,317
ATTORNEY/AGENT INFORMATION:
NAME: PARKER, DAVID L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UTSD:292/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512-320-7200
TELEFAX: 512-474-7577
TELEX: NOT APPLICABLE
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acid residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US93-06171-10
Query Match 1.8%; Score 6; DB 5; Length 19;

Best Local Similarity 100.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 LSPML 33
|||||
Db 11 LSPML 16

RESULT 30
US-08-839-624-24
Sequence 24, Application US/08839624
Patent No. 625045
GENERAL INFORMATION:
APPLICANT: KARN et al.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR COMBATING
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, Inc.
STREET: One Financial Center
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-08/839,624
FILING DATE: April 15, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/78191
FILING DATE: 15-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/017,268
FILING DATE: 13-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kathleen M. Williams
REGISTRATION NUMBER: 34,580
REFERENCE/DOCKET NUMBER: 4255/5390
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-839-624-24
Query Match 1.8%; Score 6; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 KTFAC 321
|||||
Db 20 KTFAC 25

RESULT 31
US-09-150-812-24
Sequence 24, Application US/09150812
Patent No. 639891
GENERAL INFORMATION:
APPLICANT: KARN et al.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR COMBATING
HIV INFECTION
NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:
ADDRESSER: Banner & Witcoff, Inc.
STREET: One Financial Center
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/150,812
FILING DATE: 11-Sep-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/839,624
FILING DATE: <Unknown>
APPLICATION NUMBER: US 60/017,268
FILING DATE: 13-MAY-1996

ATTORNEY/AGENT INFORMATION:
NAME: Kathleen M. Williams
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 3255/5390
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111

INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-09-150-812-24

Query Match 1.8%, Score 6, DB 4, Length 25;
Best Local Similarity 100.0%; Pred. No. 59,
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 316 KKPAC 321
11111
DB 20 KKPAC 25

RESULT 32
US-08-839-624-23
Sequence 23, Application US/08839624
Patent No. 6225045
GENERAL INFORMATION:
APPLICANT: Karn et al.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR COMBATING
NUMBER OF SEQUENCES: 43
FILING DATE: 11-Sep-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/839,624
FILING DATE: 13-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kathleen M. Williams
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 3255/5390
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/839,624
FILING DATE: April 15, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: P01/3896/78191
FILING DATE: 15-Apr-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/017,268
FILING DATE: 13-MAY-1996

ATTORNEY/AGENT INFORMATION:
NAME: Kathleen M. Williams
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 3255/5390
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111

INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-839-624-23

Query Match 1.8%, Score 6, DB 4, Length 33;
Best Local Similarity 100.0%; Pred. No. 76,
Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 316 KKPAC 321
11111
DB 28 KKPAC 33

RESULT 33
US-09-150-812-23
Sequence 23, Application US/09150812
Patent No. 6395891
GENERAL INFORMATION:
APPLICANT: Karn et al.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR COMBATING
NUMBER OF SEQUENCES: 43
FILING DATE: 11-Sep-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/839,624
FILING DATE: 13-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kathleen M. Williams
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 3255/5390
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/150,812
FILING DATE: 11-Sep-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/839,624
FILING DATE: 13-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kathleen M. Williams
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 3255/5390
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111

INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown

MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 23;
US-09-150-812-23

Query Match 1.8%; Score 6; DB 4; Length 33;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 316 KKPAC 321
DB 28 KKPAC 33

RESULT 34
US-08-839-624-25
Sequence 25, Application US/08839624
Patent No. 6225045
GENERAL INFORMATION:
APPLICANT: Kain et al.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR COMBATING
TITLE OF INVENTION: HIV INFECTION
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, Inc.
STREET: One Financial Center
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/819,624
FILING DATE: April 15, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/78191
FILING DATE: 15-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/017,268
FILING DATE: 13-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kathleen M. Williams
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 3255/5390
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
FAX: 617-345-9111
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-839-624-25

Query Match 1.8%; Score 6; DB 4; Length 37;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 316 KKPAC 321
DB 32 KKPAC 37

RESULT 35
US-09-150-812-25
Sequence 25, Application US/09150812
Patent No. 6195891

GENERAL INFORMATION:
APPLICANT: Kain et al.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR COMBATING
TITLE OF INVENTION: HIV INFECTION
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, Inc.
STREET: One Financial Center
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/150,812
FILING DATE: 11-SEP-1998
CLASSIFICATION: <UNKNOWN>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/839,624
FILING DATE: <UNKNOWN>
APPLICATION NUMBER: US 60/017,268
FILING DATE: 13-MAY-1996

ATTORNEY/AGENT INFORMATION:
NAME: Kathleen M. Williams
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 3255/5390
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
FAX: 617-345-9111

INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 25;
US-09-150-812-25

Query Match 1.8%; Score 6; DB 4; Length 37;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 316 KKPAC 321
DB 32 KKPAC 37

RESULT 36
US-08-189-331-143
Sequence 143, Application US/08189331
Patent No. 5747334

GENERAL INFORMATION:
APPLICANT: Kay, H. K.
TITLE OF INVENTION: Totally Synthetic Affinity Reagents
NUMBER OF SEQUENCES: 186
CORRESPONDENCE ADDRESS:
ADDRESSEE: Penrite & Edmunds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/189,331
FILING DATE: Concurrently herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-155
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 143:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-189-331-143

Query Match 1.8%; Score 6; DB 1; Length 39;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 RPSWVG 255
DB 6 RPSWVG 11

RESULT 37
US-08-471-068-143
Sequence 143, Application US/08471068
Patent No. 5948635
GENERAL INFORMATION:
APPLICANT: Kay, B. K.
APPLICANT: Fowlkes, R. M.
TITLE OF INVENTION: Totally Synthetic Affinity Reagents
NUMBER OF SEQUENCES: 186
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,068
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/189,331
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-155
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 143:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide

US-08-471-068-143

Query Match 1.8%; Score 6; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 250 RPSWVG 255
DB 6 RPSWVG 11

RESULT 38
US-08-689-421-1
Sequence 1, Application US/08689421
Patent No. 6008029
GENERAL INFORMATION:
APPLICANT: Yaver, Debbie S.
APPLICANT: Brown, Kimberly M.
APPLICANT: Kauppinen, Sakari
APPLICANT: Halkier, Torben P.
TITLE OF INVENTION: PURIFIED COPOLYMERS LACTASES AND NUC
TITLE OF INVENTION: ENCODING SAME
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 60080296 No. 60080296disk of No. 60080296
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/689,421
FILING DATE: 9-AUG-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Harrington, James J.
REGISTRATION NUMBER: 38,711
REFERENCE/DOCKET NUMBER: 4554,204-WO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0124
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-689-421-1

Query Match 1.8%; Score 6; DB 3; Length 40;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 309 VDNVWI 314
DB 33 VDNVWI 38

RESULT 39
US-09-389-528-1
Sequence 1, Application US/09389528
Patent No. 6207430
GENERAL INFORMATION:
APPLICANT: Yaver, Debbie S.
APPLICANT: Brown, Kimberly M.
APPLICANT: Kauppinen, Sakari
APPLICANT: Halkier, Torben P.

```

? TITLE OF INVENTION: PURIFIED COPRINUS LACCASES AND NUCLEIC ACIDS
? TITLE OF INVENTION: ENCODING SAME
? NUMBER OF SEQUENCES: 36
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: No. 6207440- No. 6207440disk of No. 6207440th America, Inc.
? STREET: 405 Lexington Avenue, 64th Floor
? CITY: New York
? STATE: New York
? COUNTRY: United States of America
? ZIP: 10174-6401
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #10, Version #130
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/389,528
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US/09/699,421
? FILING DATE: 9-AUG-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Harrington, James J.
? REGISTRATION NUMBER: 38,711
? REFERENCE/DOCKET NUMBER: 4554,204-WO
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 212-867-0133
? TELEFAX: 212-878-9655
? INFORMATION FOR SEQ ID NO. 1
? SEQUENCE CHARACTERISTICS:
? LENGTH: 40 amino acids
? TYPE: amino acid
? STRANDEDNESS:
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
US-09-389-528-1

```

```

Query Match 1.8%; Score 6; DB 4; Length 40;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 309 VDNWMI 314
DB 33 VDNWMI 38

```

```

RESULT 40
US-09-181-827A-1
? Sequence 1, Application US/09181827A
? Patent No. 6242232
? GENERAL INFORMATION:
? APPLICANT: Yaver, Debbie S.
? APPLICANT: Brown, Kimberly M.
? APPLICANT: Kaupinen, Sakari
? APPLICANT: Halkier, Torben P.
? TITLE OF INVENTION: Purified Coprinus Laccases And Nucleic
? FILE REFERENCE: 4554,200-US
? CURRENT APPLICATION NUMBER: US/09/181,827A
? CURRENT FILING DATE: 1998-10-28
? PRIOR APPLICATION NUMBER: 60/002,800
? PRIOR FILING DATE: 1995-08-25
? NUMBER OF SEQ ID NOS: 36
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 1
? LENGTH: 40
? TYPE: PRT
? ORGANISM: Coprinus cinereus
US-09-181-827A-1

```

```

Query Match 1.8%; Score 6; DB 4; Length 40;
Best Local Similarity 100.0%; Pred. No. 91;

```

```

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 309 VDNWMI 314
DB 33 VDNWMI 38

```

```

RESULT 41
US-08-839-624-26
? Sequence 26, Application US/08839624
? Patent No. 6225045
? GENERAL INFORMATION:
? APPLICANT: Karn et al.
? TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR COMBATING
? NUMBER OF SEQUENCES: 43
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Hammer & Witcoff, Inc.
? STREET: One Financial Center
? CITY: Boston
? STATE: Massachusetts
? COUNTRY: USA
? ZIP: 02111
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Wordperfect 6.1
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/839,624
? FILING DATE: April 15, 1997
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: PCT/396/78191
? FILING DATE: 15-APR-1996
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/017,268
? FILING DATE: 13-MAY-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Kathleen M. Williams
? REGISTRATION NUMBER: 34,380
? REFERENCE/DOCKET NUMBER: 3255/5390
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 617-345-9100
? TELEFAX: 617-345-9111
? INFORMATION FOR SEQ ID NO: 25:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 49 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: unknown
? MOLECULE TYPE: peptide
US-08-839-624-26

```

```

Query Match 1.8%; Score 6; DB 4; Length 49;
Best Local Similarity 100.0%; Pred. No. 1,1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 316 KRPAC 421
DB 44 KRPAC 49

```

```

RESULT 42
US-09-150-812-26
? Sequence 26, Application US/09150812
? Patent No. 6305891
? GENERAL INFORMATION:
? APPLICANT: Karn et al.
? TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR COMBATING
? HIV INFECTION
? NUMBER OF SEQUENCES: 43
? CORRESPONDENCE ADDRESS:

```


ADDRESSEE: Banner & Wilcoff, Inc.
STREET: One Financial Center
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/150,812
FILING DATE: 11-Sep-1998
CLASSIFICATION: <unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/839,624
FILING DATE: <unknown>
APPLICATION NUMBER: US 60/017,268
FILING DATE: 13-MAY-1996

ATTORNEY/AGENT INFORMATION:
NAME: Kathleen M. Williams
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 3255/5390
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111

INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-150-812-26

Query Match 1.8%; Score 6; DB 4; Length 49;
Best Local Similarity 100.0%; Pred No. 1 1e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

09 316 KKPAC 321
11111
Db 44 KKPAC 49

RESULT 44
US-08-946-165A-308
Sequence 308, Application US/08936165A
Patent No. 6348582

GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Burnham, Martin
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Lonetto, Michael
APPLICANT: Nicholas, Richard
APPLICANT: Pratt, Julie
APPLICANT: Reichard, Richard
APPLICANT: Rosenberg, Martin
APPLICANT: Ward, Judith

TITLE OF INVENTION: No. 6348582: Prokaryotic Polynucleotides,
TITLE OF INVENTION: Polypeptides and Their Uses
NUMBER OF SEQUENCES: 534
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithkline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/946,165A
FILING DATE: 24-SEP-1997
CLASSIFICATION: 536
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 60/027,872
FILING DATE: 24-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmi, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50649
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:

INFORMATION FOR SEQ ID NO: 408:
SEQUENCE CHARACTERISTICS:
LENGTH: 52 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-08-936-165A-308

Query Match 1.4%; Score 6; DB 4; Length 52;
Best Local Similarity 100.0%; Pred No. 1 2e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

07 224 EVRAOL 229
111111
Db 17 EVRAOL 22

RESULT 44
US-07-664-989B-78
Sequence 78, Application US/0764989B
Patent No. 5223409

GENERAL INFORMATION:
APPLICANT: Ladner, Robert Charles
APPLICANT: Guterman, Sonia Kosow
APPLICANT: Roberts, Bruce Lindsay
APPLICANT: Matland, William
APPLICANT: Ley, Arthur Charles
APPLICANT: Keel, Rachel Barbara
TITLE OF INVENTION: Directed Evolution of No. 5223409
TITLE OF INVENTION: Binding Proteins
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brody and Netmark
STREET: 419 Seventh Street, N.W.
CITY: Suite 300
STATE: Washington, DC
COUNTRY: USA
ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 4.2

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07664,989B
FILING DATE: 19910301
CLASSIFICATION: 530
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: P-1,765B,204731
FILING DATE: 01-SEP-1989
PRIORITY APPLICATION DATA: 07/487,008
APPLICATION NUMBER: 02-MAR-1990

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/240,160
 FILING DATE: 02-SEP-1988
 ATTORNEY/AGENT INFORMATION:
 NAME: Cooper, Iver P.
 REGISTRATION NUMBER: 28005
 REFERENCE/DOCKET NUMBER: LADNER 7
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-628-5197
 TELEFAX: 202-737-3528
 INFORMATION FOR SEQ ID NO: 78:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 58 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-07-664-9898-78

Query Match 1.8%; Score 6; DB 1; Length 58;
 Best Local Similarity 100.0%; Pred. No. 1,3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 GCLCHG 21
 |||||
 DB 37 GCLCHG 42

RESULT 45
 US-08-358-160-55
 Sequence 55, Application US/08358160
 Patent No. 5663143
 GENERAL INFORMATION:
 APPLICANT: LEY, Arthur C.
 APPLICANT: LADNER, Robert C.
 APPLICANT: GUTERMAN, Sonia K.
 APPLICANT: ROBERTS, Bruce L.
 APPLICANT: MARKLAND, William
 APPLICANT: KENT, Rachel H.
 TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ
 TITLE OF INVENTION: DOMAINS THAT INHIBIT HUMAN NEUTROPHIL ELASTASE
 NUMBER OF SEQUENCES: 234
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BROMDY AND NEIMARK
 STREET: 419 Seventh Street, N.W. Suite 300
 CITY: Washington
 STATE: District of Columbia
 COUNTRY: USA
 ZIP: 20004
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/358,160
 FILING DATE: 16-DEC-1994
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/143,031
 FILING DATE: 13-OCT-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/009,319
 FILING DATE: 26-JAN-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/664,989
 FILING DATE: 01-MAR-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/487,063
 FILING DATE: 02-MAR-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/240,160
 FILING DATE: 02-SEP-1988
 ATTORNEY/AGENT INFORMATION:

NAME: Cooper, Iver P.
 REGISTRATION NUMBER: 28,005
 REFERENCE/DOCKET NUMBER: LEY-1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-628-5197
 TELEFAX: 202-737-3528
 TELEEX: 248633
 INFORMATION FOR SEQ ID NO: 55:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 58 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-358-160-55

Query Match 1.8%; Score 6; DB 1; Length 58;
 Best Local Similarity 100.0%; Pred. No. 1,3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 GCLCHG 21
 |||||
 DB 37 GCLCHG 42

Search completed: December 7, 2002, 11:18:14
 Job time: 22 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: December 7, 2002, 11:15:38, Search time 13 seconds

(without alignments)
406,058 Million cell updates/sec

Title: US-09-831-458A-12

Perfect score: 125

Sequence: 1 MSMSKPRVVOGLHGLGH PCGVNMYWCKRPAACEFDE 325

Scoring table:

Gapop 60 0, Gapext 60 0

Searched: 103943 seqs, 16242309 residues

Word size: 5

Total number of hits satisfying chosen parameters: 3135

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Database: Published Applications-AA*

1: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/prodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	15.4	50	10	US-09-864-761-40945
2	19	2.8	31	9	US-10-125-452-27
3	9	2.8	339	10	US-09-983-531A-26
4	9	2.8	427	10	US-09-983-531A-16
5	9	2.8	427	10	US-09-983-531A-20
6	9	2.8	778	9	US-10-125-470-15
7	9	2.8	778	9	US-10-125-452-16
8	9	2.8	918	9	US-10-020-733-4
9	9	2.8	918	9	US-10-020-733-4
10	9	2.8	918	9	US-10-125-470-9
11	9	2.8	926	9	US-10-125-452-9
12	9	2.8	956	9	US-10-020-733-2
13	9	2.8	963	9	US-10-020-733-8
14	8	2.5	472	10	US-09-864-582-20
15	8	2.2	144	10	US-09-748-673-427
16	7	2.2	109	10	US-09-825-882-8
17	7	2.2	367	10	US-09-815-242-10118
18	7	2.2	456	10	US-09-815-242-14020
19	7	2.2	548	10	US-09-732-350-4

20	7	2.2	720	10	US-09-919-497-83	Appl 1
21	7	2.2	769	9	US-10-072-841-31	Appl 1
22	6	1.8	29	10	US-09-942-172-15	Appl 1
23	6	1.8	29	10	US-09-982-172-215	Appl 1
24	6	1.8	31	10	US-09-864-761-48338	Appl 1
25	6	1.8	37	10	US-09-864-761-39821	Appl 1
26	6	1.8	38	10	US-09-864-761-48289	Appl 1
27	6	1.8	44	10	US-09-864-761-44451	Appl 1
28	6	1.8	49	10	US-09-864-761-38145	Appl 1
29	6	1.8	51	10	US-09-864-761-40929	Appl 1
30	6	1.8	52	10	US-09-949-980-308	Appl 1
31	6	1.8	54	10	US-09-764-860-353	Appl 1
32	6	1.8	57	10	US-09-764-877-1727	Appl 1
33	6	1.8	58	10	US-09-781-988-78	Appl 1
34	6	1.8	90	10	US-09-867-550-1392	Appl 1
35	6	1.8	102	9	US-10-004-381-17	Appl 1
36	6	1.8	111	10	US-09-764-854-786	Appl 1
37	6	1.8	117	10	US-09-864-761-45806	Appl 1
38	6	1.8	119	10	US-09-764-870-402	Appl 1
39	6	1.8	131	10	US-09-815-242-12092	Appl 1
40	6	1.8	138	10	US-09-216-393-66	Appl 1
41	6	1.8	139	10	US-09-815-242-10702	Appl 1
42	6	1.8	153	9	US-09-764-868-1057	Appl 1
43	6	1.8	154	10	US-09-825-401-27	Appl 1
44	6	1.8	154	10	US-09-604-2878-47	Appl 1
45	6	1.8	154	10	US-10-007-805-471	Appl 1

ALIGNMENTS

RESULT 1
US-09-864-761-40945
Sequence 40945, Application US-09-864-761
Patent No. US-2004/08763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GINOM-DERIVED SINGLE EXON NUCLEOTIDE SEQUENCES
FILE REFERENCE: Aemica-X-1
CURRENT FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: US-60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US-60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US-60/264,306
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: GR-24264.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US-60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US-60/200,120
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30

```

PRIORITY APPLICATION NUMBER: PCT/US01/00670
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: US 60/234,687
PRIORITY FILING DATE: 2000-09-21
PRIORITY APPLICATION NUMBER: US 09/608,408
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: US 09/774,203
PRIORITY FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 40945
LENGTH: 50
TYPE: PRI
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC008812.4
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
OTHER INFORMATION: SWISSPROT Hit: P26717, EVALUE 6.09e-07
OTHER INFORMATION: EST_HUMAN Hit: H00958.1, EVALUE 2.00e-25
US-09-864-761-40945

```

```

Query Match: 15.4%; Score 9; DB 10; Length 50;
Best Local Similarity 100.0%; Pred. No. 2,80-30;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 184 PTCHEKMTFFGNCYEMNSQPMHDSVACGEVRAQVVIKTAEEQ 238
DB 1 RICHCHCKDKTFFGNCYEMNSQPMHDSVACGEVRAQVVIKTAEEQ 50

```

```

RESULT 2
US-10-125-452-27
Sequence 27; Application US/20125452
Patent No. US20020173640A1
GENERAL INFORMATION:
APPLICANT: Ruden et al.
TITLE OF INVENTION: ADAM Polynucleotides, Polypeptides, and Antibodies
FILE REFERENCE: p700bnp2
CURRENT APPLICATION NUMBER: US/10/125,452
PRIORITY FILING DATE: 2002-04-19
PRIORITY APPLICATION NUMBER: 09/955,504
PRIORITY FILING DATE: 2001-09-19
PRIORITY APPLICATION NUMBER: 09/712,907
PRIORITY FILING DATE: 2000-11-16
PRIORITY APPLICATION NUMBER: P21/99500/14308
PRIORITY FILING DATE: 2000-05-25
PRIORITY APPLICATION NUMBER: 60/178,717
PRIORITY FILING DATE: 2000-01-28
PRIORITY APPLICATION NUMBER: 60/142,930
PRIORITY FILING DATE: 1999-07-09
PRIORITY APPLICATION NUMBER: 60/136,388
PRIORITY FILING DATE: 1999-05-27
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 27
LENGTH: 41
TYPE: PRI
ORGANISM: Homo sapiens
US-10-125-452-27

```

```

Query Match: 2.8%; Score 9; DB 9; Length 31;
Best Local Similarity 100.0%; Pred. No. 0 083;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 34 AGVAVALLV 42
DB 1 AGVAVALLV 11

```

```

RESULT 3
US-09-983-531A-16
Sequence 26; Application US/09983531A
Patent No. US20020147132A1
GENERAL INFORMATION:
APPLICANT: Fujisawa, Atsuko
APPLICANT: Yamakawa, Toru
APPLICANT: Shitakawa, Kanon
APPLICANT: Chitose, Ofii
APPLICANT: Ogawa, Naoki
TITLE OF INVENTION: Meltrin
FILE REFERENCE: 11-22-99 sequence submission
CURRENT APPLICATION NUMBER: US/09/983,531A
PRIORITY FILING DATE: 2001-10-24
PRIORITY APPLICATION NUMBER: JP 8-61756
PRIORITY FILING DATE: 1996-02-24
PRIORITY APPLICATION NUMBER: PTC/JP96/703017
PRIORITY FILING DATE: 1996-10-17
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 26
LENGTH: 349
TYPE: PRI
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Other
OTHER INFORMATION: JMI094(PM61 beta 24C); human meltrin beta
US-09-983-531A-16

```

```

Query Match: 2.8%; Score 9; DB 10; Length 349;
Best Local Similarity 100.0%; Pred. No. 0 67;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 34 AGVAVALLV 42
DB 245 AGVAVALLV 253

```

```

RESULT 4
US-09-983-531A-16
Sequence 16; Application US/09983531A
Patent No. US20020147132A1
GENERAL INFORMATION:
APPLICANT: Fujisawa, Atsuko
APPLICANT: Yamakawa, Toru
APPLICANT: Shitakawa, Kanon
APPLICANT: Chitose, Ofii
APPLICANT: Ogawa, Naoki
TITLE OF INVENTION: Meltrin
FILE REFERENCE: 11-22-99 sequence submission
CURRENT APPLICATION NUMBER: US/09/983,531A
PRIORITY FILING DATE: 2001-10-24
PRIORITY APPLICATION NUMBER: JP 8-61756
PRIORITY FILING DATE: 1996-02-24
PRIORITY APPLICATION NUMBER: PTC/JP96/703017
PRIORITY FILING DATE: 1996-10-17
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 16
LENGTH: 394
TYPE: PRI
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Human meltrin
OTHER INFORMATION: beta derived from cDNA
US-09-983-531A-16

```

```

Query Match: 2.8%; Score 9; DB 10; Length 394;
Best Local Similarity 100.0%; Pred. No. 0 76;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 34 AGVLVAILV 42
|||||
Db 333 AGVLVAILV 341

RESULT 5

US-09-983-531A-20
Sequence 20, Application US/09983531A
Patent No. US20020147132A1
GENERAL INFORMATION:
APPLICANT: Fujisawa, Atsuko
APPLICANT: Yamakawa, Toru
APPLICANT: Shirakawa, Kamon
APPLICANT: Chitose, Oriti
APPLICANT: Ogawa, Naoki
TITLE OF INVENTION: Meltrins
FILE REFERENCE: 11-22-99 sequence submission
CURRENT APPLICATION NUMBER: US/09/983,531A
CURRENT FILING DATE: 2001-10-24
PRIOR APPLICATION NUMBER: JP 8-61756
PRIOR FILING DATE: 1996-02-23
PRIOR APPLICATION NUMBER: PCT/JP96/03017
PRIOR FILING DATE: 1996-10-17
NUMBER OF SEQ ID NOS: 44
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 20
LENGTH: 427
TYPE: PRT
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Clones:
OTHER INFORMATION: JM109(pmel-beta-24C) and JM109(pmel-beta-24N),
US-09-983-531A-20

Query Match: 2.8%; Score 9, DB 10, Length 427;
Best Local Similarity 100.0%; Pred. No. 0.81;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 34 AGVLVAILV 42
|||||
Db 333 AGVLVAILV 341

RESULT 6

US-10-125-470-16
Sequence 16, Application US/10125470
Patent No. US20020165377A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: ADAM Polynucleotides, Polypeptides, and Antibodies
FILE REFERENCE: PTO06P1
CURRENT APPLICATION NUMBER: US/10/125,470
CURRENT FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: US/09/712,907A
PRIOR FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: 50/130,588
PRIOR FILING DATE: 1999-05-27
PRIOR APPLICATION NUMBER: 60/142,930
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 60/178,717
PRIOR FILING DATE: 2000-01-28
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 16
LENGTH: 778
TYPE: PRT
ORGANISM: Homo sapiens
US-10-125-470-16

Query Match: 2.8%; Score 9, DB 9, Length 778;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 34 AGVLVAILV 42
|||||
Db 705 AGVLVAILV 713

RESULT 7

US-10-125-452-16
Sequence 16, Application US/10125472
Patent No. US20020173640A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: ADAM Polynucleotides, Polypeptides, and
FILE REFERENCE: PTO06P2
CURRENT APPLICATION NUMBER: US/10/125,452
CURRENT FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/955,504
PRIOR FILING DATE: 2001-09-19
PRIOR APPLICATION NUMBER: 09/712,607
PRIOR FILING DATE: 2000-11-16
PRIOR APPLICATION NUMBER: PCT/US00/14408
PRIOR FILING DATE: 2000-05-25
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: 60/142,730
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 60/161,088
PRIOR FILING DATE: 1999-05-27
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 16
LENGTH: 778
TYPE: PRT
ORGANISM: Homo sapiens
US-10-125-452-16

Query Match: 2.8%; Score 9, DB 9, Length 778;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0;

OY 34 AGVLVAILV 42
|||||
Db 705 AGVLVAILV 713

RESULT 8

US-10-020-733-4
Sequence 4, Application US/10020733
Patent No. US20020161214A1
GENERAL INFORMATION:
APPLICANT: Walke, D. Wade
APPLICANT: Wilgowski, Nathaniel L.
TITLE OF INVENTION: Fridde, Carl Johan
FILE REFERENCE: LEX-0263-058A
CURRENT APPLICATION NUMBER: US/10/020,733
CURRENT FILING DATE: 2001-10-30
PRIOR APPLICATION NUMBER: US 60/244,999
PRIOR FILING DATE: 2000-11-01
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 918
TYPE: PRT
ORGANISM: homo sapiens
US-10-020-733-4

Query Match: 2.8%; Score 9, DB 9, Length 918;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0;

OY 34 AGVLVAILV 42
|||||

DB 705 ACVVALIV 713

RESULT 9

US-10-125-470-9
 : Sequence 9, Application US/10125470
 : Patent No. US20020165377A1
 : GENERAL INFORMATION:
 : APPLICANT: Ribben et al.
 : TITLE OF INVENTION: ADAM Polymorphonuclear Leukocytes, Polypeptides, and Antibodies
 : FILE REFERENCE: PT006P1
 : CURRENT APPLICATION NUMBER: US/10/125,470
 : PRIOR FILING DATE: 2000-04-19
 : PRIOR APPLICATION NUMBER: US/09/712,907A
 : PRIOR FILING DATE: 2001-03-07
 : PRIOR APPLICATION NUMBER: 60/136,388
 : PRIOR FILING DATE: 1999-06-27
 : PRIOR APPLICATION NUMBER: 60/142,930
 : PRIOR FILING DATE: 1999-07-09
 : PRIOR APPLICATION NUMBER: 60/178,717
 : PRIOR FILING DATE: 2000-01-28
 : NUMBER OF SEQ ID NOS: 24
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 9
 : LENGTH: 918
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 US-10-125-470-9

Query Match

2.8% Score 9; DB 9; Length 918;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 ACVVALIV 42

DB 705 ACVVALIV 713

RESULT 10

US-10-125-452-9
 : Sequence 9, Application US/10125452
 : Patent No. US20020173640A1
 : GENERAL INFORMATION:
 : APPLICANT: Ribben et al.
 : TITLE OF INVENTION: ADAM Polymorphonuclear Leukocytes, Polypeptides, and Antibodies
 : FILE REFERENCE: PT006P2
 : CURRENT APPLICATION NUMBER: US/10/125,452
 : PRIOR FILING DATE: 2002-04-19
 : PRIOR APPLICATION NUMBER: 09/055,504
 : PRIOR FILING DATE: 2001-09-19
 : PRIOR APPLICATION NUMBER: 09/712,907
 : PRIOR FILING DATE: 2000-11-16
 : PRIOR APPLICATION NUMBER: PCT/US00/14108
 : PRIOR FILING DATE: 2000-05-25
 : PRIOR APPLICATION NUMBER: 60/178,717
 : PRIOR FILING DATE: 2000-01-28
 : PRIOR APPLICATION NUMBER: 60/142,930
 : PRIOR FILING DATE: 1999-07-09
 : PRIOR APPLICATION NUMBER: 60/136,388
 : PRIOR FILING DATE: 1999-06-27
 : NUMBER OF SEQ ID NOS: 38
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 9
 : LENGTH: 918
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 US-10-125-452-9

Query Match

2.8% Score 9; DB 9; Length 918;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 ACVVALIV 42

DB 705 ACVVALIV 713

RESULT 11

US-10-020-733-2
 : Sequence 2, Application US/10020733
 : Patent No. US20020161214A1
 : GENERAL INFORMATION:
 : APPLICANT: Waite, D. Wade
 : APPLICANT: Wilgowski, Nathaniel L.
 : TITLE OF INVENTION: No. US20020161214A1; Human Proteases and Polymorphonuclear Leukocytes
 : FILE REFERENCE: LEX-0263-USA
 : CURRENT APPLICATION NUMBER: US/10/020,733
 : PRIOR FILING DATE: 2001-10-10
 : PRIOR APPLICATION NUMBER: US 60/244,939
 : PRIOR FILING DATE: 2000-11-01
 : NUMBER OF SEQ ID NOS: 9
 : SOFTWARE: FASTSEQ for Windows Version 4.0
 : SEQ ID NO 2
 : LENGTH: 926
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 US-10-020-733-2

Query Match

2.8% Score 5; DB 9; Length 926;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 ACVVALIV 42

DB 705 ACVVALIV 713

RESULT 12

US-10-020-733-8
 : Sequence 8, Application US/10020733
 : Patent No. US20020161214A1
 : GENERAL INFORMATION:
 : APPLICANT: Waite, D. Wade
 : APPLICANT: Wilgowski, Nathaniel L.
 : TITLE OF INVENTION: No. US20020161214A1; Human Proteases and Polymorphonuclear Leukocytes
 : FILE REFERENCE: LEX-0263-USA
 : CURRENT APPLICATION NUMBER: US/10/020,733
 : PRIOR FILING DATE: 2001-10-10
 : PRIOR APPLICATION NUMBER: US 60/244,939
 : PRIOR FILING DATE: 2000-11-01
 : NUMBER OF SEQ ID NOS: 9
 : SOFTWARE: FASTSEQ for Windows Version 4.0
 : SEQ ID NO 8
 : LENGTH: 955
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 US-10-020-733-8

Query Match

2.8% Score 9; DB 9; Length 955;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 ACVVALIV 42

DB 705 ACVVALIV 713

RESULT 13

US-10-020-733-6
 : Sequence 6, Application US/10020733
 : Patent No. US20020161214A1
 : GENERAL INFORMATION:
 : APPLICANT: Waite, D. Wade
 : APPLICANT: Wilgowski, Nathaniel L.

```

: APPLICANT: Fiddle, Carl Johan
: TITLE OF INVENTION: NO. US20020161214A1e1 Human Proteases and Polymorphic Iles Encoding
: FILE REFERENCE: LEX-0263-085A
: CURRENT APPLICATION NUMBER: US/10/020,733
: CURRENT FILING DATE: 2001-10-30
: PRIOR APPLICATION NUMBER: US 60/244,939
: PRIOR FILING DATE: 2000-11-01
: NUMBER OF SEQ ID NOS: 9
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 6
: LENGTH: 963
: TYPE: PRT
: ORGANISM: homo sapiens
: US-10-020-733-6

Query Match
Best Local Similarity 100.0%; Score 9; DB 9; Length 963;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 AGVLAIVL 42
DB 705 AGVLAIVL 713

```

```

RESULT 14
US-09-866-582-29
: Sequence 29, Application US/09866582
: Patent No. US20020127620A1
: GENERAL INFORMATION:
: APPLICANT: Witman, George R.
: APPLICANT: Pazour, Gregory J.
: APPLICANT: Rosenbaum, Joel L.
: APPLICANT: Cole, Douglas G.
: TITLE OF INVENTION: INTRACELLULAR TRANSPORT
: FILE REFERENCE: 07917-145001
: CURRENT APPLICATION NUMBER: US/09/866,582
: CURRENT FILING DATE: 2001-05-24
: PRIOR APPLICATION NUMBER: US 60/206,923
: PRIOR FILING DATE: 2000-05-24
: NUMBER OF SEQ ID NOS: 50
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 29
: LENGTH: 472
: TYPE: PRT
: ORGANISM: Caenorhabditis elegans
: US-09-866-582-29

Query Match
Best Local Similarity 100.0%; Score 8; DB 10; Length 472;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 DSKKQKQI 176
DB 21 DSKKQKQI 28

```

```

RESULT 15
US-09-738-973-427
: Sequence 427, Application US/09738973
: Patent No. US20020110563A1
: GENERAL INFORMATION:
: APPLICANT: Reed, Steven G.
: APPLICANT: Henderson, Robert A.
: APPLICANT: Lodes, Michael J.
: APPLICANT: Fling, Steven P.
: APPLICANT: Mohamath, Radoch
: APPLICANT: Algate, Paul A.
: APPLICANT: Secrist, Heather
: APPLICANT: Indrias, Carol Joseph
: APPLICANT: Renson, Darin K.
: APPLICANT: Elliot, Mark
: APPLICANT: Mannion, Jaee
: APPLICANT: Kalos, Michael D.

```

```

: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
: TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CAN
: FILE REFERENCE: 210121.47509
: CURRENT APPLICATION NUMBER: US/09/738,973
: CURRENT FILING DATE: 2000-12-14
: NUMBER OF SEQ ID NOS: 587
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 427
: LENGTH: 184
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-738-973-427

Query Match
Best Local Similarity 100.0%; Score 7; DB 10; Length 18
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 223 QEVRAQL 229
DB 162 QEVRAQL 168

```

```

RESULT 16
US-09-825-882-8
: Sequence 8, Application US/09825882
: Patent No. US20020094551A1
: GENERAL INFORMATION:
: APPLICANT: ADLER, JON ELLIOT
: TITLE OF INVENTION: T2R TASTE RECEPTORS AND GENES ENCODING
: FILE REFERENCE: 078063/0274527X1
: CURRENT APPLICATION NUMBER: US/09/825,882
: CURRENT FILING DATE: 2001-04-05
: PRIOR APPLICATION NUMBER: 60/195,532
: PRIOR FILING DATE: 2000-04-07
: PRIOR APPLICATION NUMBER: 60/247,014
: PRIOR FILING DATE: 2000-11-13
: NUMBER OF SEQ ID NOS: 31
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 8
: LENGTH: 409
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-825-882-8

Query Match
Best Local Similarity 100.0%; Score 7; DB 10; Length 4
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 LLSFML 33
DB 192 LLSFML 198

```

```

RESULT 17
US-09-815-242-10118
: Sequence 10118, Application US/09815242
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Frawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert L.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes in
: FILE REFERENCE: ELITRA.011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: CURRENT FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21

```

```

1 PRIOR APPLICATION NUMBER: 60/257,848
2 PRIOR FILING DATE: 2000-05-23
3 PRIOR APPLICATION NUMBER: 60/207,727
4 PRIOR FILING DATE: 2000-05-26
5 PRIOR APPLICATION NUMBER: 60/242,578
6 PRIOR FILING DATE: 2000-10-23
7 PRIOR APPLICATION NUMBER: 60/253,625
8 PRIOR FILING DATE: 2000-11-27
9 PRIOR APPLICATION NUMBER: 60/257,931
10 PRIOR FILING DATE: 2000-12-22
11 PRIOR APPLICATION NUMBER: 60/269,308
12 PRIOR FILING DATE: 2001-02-16
13 NUMBER OF SEQ ID NOS: 14110
14 SOFTWARE: FASTSEQ for Windows Version 4.0
15 SEQ ID NO 10118
16 LENGTH: 467
17 TYPE: PRT
18 ORGANISM: Escherichia coli
19 US-09-815-242-10118

```

```

Query Match
Best Local Similarity 2.2%, Score 7, DB 10, Length 367,
Matches 7, Conservative 0, Mismatches 0, Indels 0, Gaps 0

```

```

CY 23 LVLS: 29
DB 95 LVLS: 101

```

```

RESULT 18
US-09-815-242-14020
Sequence 14020, Application US/09/815,42
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlson, Karl L.
APPLICANT: Zyskind, Judith W
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA, 011A
CURRENT APPLICATION NUMBER: US/09/815,342
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 14020
LENGTH: 455
TYPE: PRT
ORGANISM: Salmonella typhi
US-09-815-242-14020

```

```

Query Match
Best Local Similarity 2.2%, Score 7, DB 10, Length 455,
Matches 7, Conservative 0, Mismatches 0, Indels 0, Gaps 0

```

```

CY 23 LVLS: 29
DB 183 LVLS: 189

```

```

RESULT 19
US-09-732-350-4
Sequence 4, Application US/09/732,350
Patent No. US20010041490A1
GENERAL INFORMATION:
APPLICANT: Swenson, Allan
APPLICANT: Xu, Feng
TITLE OF INVENTION: LACCASE MUTANTS
NUMBER OF SEQUENCES: 10
CORRESPONDING ADDRESS:
ADDRESSEE: No. US20010031190A1a No. US20010031490A1dsk 91 No. US20010031490A
SHEET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER SEPARABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/732,350
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 67/642,315
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Kozick, Carol
REGISTRATION NUMBER: 36,993
REFERENCE/INVENT NUMBER: 6200, 200-05
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-732-350-4

```

```

Query Match
Best Local Similarity 2.2%, Score 7, DB 10, Length 548,
Matches 7, Conservative 0, Mismatches 0, Indels 0, Gaps 0

```

```

CY 308 LVLS: 114
DB 275 LVLS: 281

```

```

RESULT 20
US-09-919-497-83
Sequence 83, Application US/09/919,497
Patent No. US2002010662A1
GENERAL INFORMATION:
APPLICANT: Muller, George L.
TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
FILE REFERENCE: B0801/7225
CURRENT APPLICATION NUMBER: US/09/919,497
CURRENT FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: US/09/221,735
PRIOR FILING DATE: 2000-07-31
NUMBER OF SEQ ID NOS: 100
SOFTWARE: Patcom in version 3.0
SEQ ID NO 83

```


LENGTH: 720
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-919-497-83

Query Match 2.28: Score 7: PR 10: Length 720:
 Best Local Similarity 100.0%: Pred. No. 88:
 Matches 7: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 269 DCSPLSP 275
 |||||
 DB 247 DCSPLSP 243

RESULT 21
 US-10-072-841-31
 Sequence 31, Application US/10072841
 Patent No. US20020164708A1
 GENERAL INFORMATION:

APPLICANT: Sheppard, Dean
 Quaranta, Vito
 Pytela, Robert
 TITLE OF INVENTION: A No. US20020164708A1: Integrin Beta Subunit and Uses Thereof
 NUMBER OF SEQUENCES: 43
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Priority, Schroeder, Brueggemann & Clark
 STREET: 4470 La Jolla Village Drive, Suite 700
 CITY: San Diego
 STATE: California
 COUNTRY: United States of America
 ZIP: 92132

COMPUTER READABLE FORM:
 MEDIUM TYPE: 5 1/4 floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1 0, Version #1 26
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/072,841
 FILING DATE: 06-Feb-2002
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/728,215
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Campbell, Cathryn A.
 REGISTRATION NUMBER: 31,815
 REFERENCE/DOCKET NUMBER: P31 8717
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 535-9001
 TELEFAX: (619) 535-8949
 INFORMATION FOR SEQ ID NO: 31:

SEQUENCE CHARACTERISTICS:
 LENGTH: 769 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 SEQUENCE DESCRIPTION: SEQ ID NO: 31:
 US-10-072-841-31

Query Match 2.28: Score 7: DB 9: Length 769:
 Best Local Similarity 100.0%: Pred. No. 93:
 Matches 7: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 71 AVEGESE 77
 |||||
 DB 341 AVEGESE 347

RESULT 22
 US-09-982-172-16
 Sequence 16, Application US/09982172
 Patent No. US20020137119A1

GENERAL INFORMATION:
 APPLICANT: Emil Israel Katz
 TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES
 TITLE OF INVENTION: DIRECTED THERAPEUTICS, AND METHODS,
 TITLE OF INVENTION: UTILIZING EACH
 FILE REFERENCE: 01/22283
 CURRENT APPLICATION NUMBER: 09/09982172
 CURRENT FILING DATE: 2001-10-19
 NUMBER OF SEQ ID NOS: 253
 SOFTWARE: Patent In version 4.1
 SEQ ID NO: 16
 LENGTH: 29
 TYPE: PRT
 ORGANISM: Artificial sequence
 FEATURE:
 OTHER INFORMATION: Computer generated synthetic peptide
 US-09-982-172-16

Query Match 1.88: Score 6: DB 10: Length 29:
 Best Local Similarity 100.0%: Pred. No. 44:
 Matches 6: Conservative 0: Mismatches 0: Indels 0:

QY 34 AGVLVA 39
 |||||
 DB 10 AGVLVA 15

RESULT 23
 US-09-982-172-215
 Sequence 215, Application US/09982172
 Patent No. US20020137119A1
 GENERAL INFORMATION:

APPLICANT: Emil Israel Katz
 TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES
 TITLE OF INVENTION: DIRECTED THERAPEUTICS, AND METHODS,
 TITLE OF INVENTION: UTILIZING EACH
 FILE REFERENCE: 01/22283
 CURRENT APPLICATION NUMBER: 09/09982172
 CURRENT FILING DATE: 2001-10-19
 NUMBER OF SEQ ID NOS: 253
 SOFTWARE: Patent In version 4.1
 SEQ ID NO: 215
 LENGTH: 29
 TYPE: PRT
 ORGANISM: Artificial sequence
 FEATURE:
 OTHER INFORMATION: Computer generated synthetic peptide
 US-09-982-172-215

Query Match 1.88: Score 6: DB 10: Length 29:
 Best Local Similarity 100.0%: Pred. No. 44:
 Matches 6: Conservative 0: Mismatches 0: Indels 0:

QY 34 AGVLVA 39
 |||||
 DB 10 AGVLVA 15

RESULT 24
 US-09-864-761-48338
 Sequence 48338, Application US/09984761
 Patent No. US20020048763A1
 GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.
 APPLICANT: Bank, David R.
 APPLICANT: Hanzel, David K.
 APPLICANT: Chen, Wensheng
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEOTIDE SEQUENCES AND METHODS FOR ANALYZING AND IDENTIFYING
 TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAYS
 FILE REFERENCE: ACOMICA-X-1
 CURRENT APPLICATION NUMBER: US/09984761
 CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US-09-072841-31

```

PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: US 60/287,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/642,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GR 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00673
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 48338
LENGTH: 31
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL121923.7
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.3
OTHER INFORMATION: EST_HUMAN HIT: A1298543.1, EVALUATE 4.006 01
US-09-864-761-48338

```

```

Query Match          1.8%; Score 6; DB 10; Length 31;
Best Local Similarity 100.0%; Pred No 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 231 VIKTAE 236
|||||
DB 15 VIKTAE 20

RESULT 25
US-09-864-761-39821
Sequence 39821, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David P.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NOCLERIC ACID PROBES USEFUL FOR
FILE REFERENCE: Acemica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456

```

```

PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/542,350
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GR 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00673
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 39821
LENGTH: 37
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004125.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.4
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.5
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.6
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.7
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.6
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.9
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.9
US-09-864-761-39821

```

```

Query Match          1.8%; Score 6; DB 10; Length 37;
Best Local Similarity 100.0%; Pred No 55;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 12 LGIAGT 17
|||||
DB 24 LGIAGT 29

RESULT 26
US-09-864-761-48289
Sequence 48289, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David P.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NOCLERIC ACID PROBES USEFUL FOR
FILE REFERENCE: Acemica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761

```

```

: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263,6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/246,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 48289
: LENGTH: 38
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC019248.3
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.4
: OTHER INFORMATION: SWISSPROT HIT: Q15404, EVALUE 2.00e-15
: OTHER INFORMATION: EST_HUMAN HIT: BF65928.1, EVALUE 3.00e-14
US-09-864-761-48289
Query Match 1.8%, Score 6, DB 10, Length 38;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 87 ERTOLK 92
DB 17 ERTOLK 22

```

```

: PRIOR APPLICATION NUMBER: US 09/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263,6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/246,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 4461
: LENGTH: 44
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC06477.2
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.94
: OTHER INFORMATION: EST_HUMAN HIT: AA179279.1, EVALUE 3.1
: OTHER INFORMATION: SWISSPROT HIT: P38613, EVALUE 2.60e-14
US-09-864-761-4461
Query Match 1.8%, Score 6, DB 10, Length 44;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 46 KVPSSL 51
DB 5 KVPSSL 10

```

```

: RESULT 27
: US-09-864-761-44461
: Sequence 44461, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: FILE REFERENCE: Acomica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23

```

```

: RESULT 28
: US-09-864-761-38145
: Sequence 38145, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUC

```

```

: TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
: FILE REFERENCE: Acomica-X-1
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,412
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/532,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: US 24263,6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00660
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Anomax Sequence Listing Engine v0rs 1.1
: SEQ ID NO: 36145
: LENGTH: 49
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO A1031026.1
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.1
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.5
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.8
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.1
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.2
: OTHER INFORMATION: EXPRESSED IN HL100, SIGNAL = 4.1
: OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.9
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.9
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.4
: US-09-864-761-38145
:
: Query Match: 1.88; Score 6; DB 10; Length 49;
: Best Local Similarity 100.0%; Pred No 70;
: Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
:
: QY 271 SPLSPS 276
: |11111
: DB 38 SPLSPS 43

```

```

: GENERAL INFORMATION:
: APPLICANT: Pharm, Shattou G.
: APPLICANT: Panz, David R.
: APPLICANT: Hanzel, David R.
: APPLICANT: Chen, Wen-sheng
: TITLE OF INVENTION: HUMAN GENEOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: FILE REFERENCE: Acomica-X-1
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/532,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: US 24263,6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Anomax Sequence Listing Engine v0rs 1.1
: SEQ ID NO: 40929
: LENGTH: 51
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO A1121772.12
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
: OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
: OTHER INFORMATION: EST_HUMAN HIT: AA360017.1, EVALUATE 4.00e+00
: OTHER INFORMATION: EST_HUMAN HIT: AA360017.1, EVALUATE 4.00e+00
: US-09-864-761-40929
:
: Query Match: 1.88; Score 6; DB 10; Length 51;
: Best Local Similarity 100.0%; Pred No. 73;
: Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
:
: QY 51 ISOBOS 56
: |11111
: DB 44 ISOBOS 49

```

```

RESULT 30
US-09-939-980-108
: Sequence 308, Application US/09939980
: Patent No. US2002008234A1
GENERAL INFORMATION:
: APPLICANT: Black, Michael
:      Burnham, Martin
:      Hodgson, John
:      Knowles, David
:      Lonetto, Michael
:      Nicholas, Richard
:      Pratt, Julie
:      Reichard, Richard
:      Rosenberg, Martin
:      Ward, Judith
TITLE OF INVENTION: No. US2002008234A1 Polynucleotide Polypeptides,
:      Polypeptides and Their Uses
NUMBER OF SEQUENCES: 534
CORRESPONDENCE ADDRESS:
: ADDRESSEE: Smithkline Beecham Corporation
: STREET: 709 Swedeland Road
: CITY: King of Prussia
: STATE: PA
: COUNTRY: USA
: ZIP: 19406-0939
COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/939,980
: FILING DATE: 27-Aug-2001
: CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/936,165
: FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
: NAME: Gimmil, Edward R
: REGISTRATION NUMBER: 38,891
: REFERENCE/DOCKET NUMBER: P50549
TELECOMMUNICATION INFORMATION:
: TELEPHONE: 610-270-4478
: TELEFAX: 610-270-5090
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 308:
: SEQUENCE CHARACTERISTICS:
:   LENGTH: 52 amino acids
:   TYPE: amino acid
:   STRANDEDNESS: single
:   TOPOLOGY: linear
MOLECULE TYPE: Protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 308:
US-09-939-980-308

Query Match          1.8% Score 6; DB 10; Length 52;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 224 EVRAQL 229
DB 17 EVRAQL 22

RESULT 31
US-09-764-860-353
: Sequence 353, Application US/09764860
: Patent No. US20020094953A1
GENERAL INFORMATION:
: APPLICANT: Rosen et al
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PC008

```

```

CURRENT APPLICATION NUMBER: US/09/764,800
: CURRENT FILING DATE: 2001-01-17
: Prior application data removed
: NUMBER OF SEQ ID NOS: 1198
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 353
: LENGTH: 54
: TYPE: PRT
: ORGANISM: Homo sapiens
FEATURE:
: NAME/KEY: SITE
: LOCATION: (12)
: OTHER INFORMATION: Xaa equals any of the naturally occur
: NAME/KEY: SITE
: LOCATION: (47)
: OTHER INFORMATION: Xaa equals any of the naturally occur
: NAME/KEY: SITE
: LOCATION: (52)
: OTHER INFORMATION: Xaa equals any of the naturally occur
US-09-764-860-353

Query Match          1.8% Score 6; DB 10; Length 5
Best Local Similarity 100.0%; Pred. No. 76;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 LAGVIV 38
DB 38 LAGVIV 43

RESULT 32
US-09-764-877-1727
: Sequence 1727, Application US/09764877
: Patent No. US20020147140A1
GENERAL INFORMATION:
: APPLICANT: Rosen et al
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PC005
CURRENT APPLICATION NUMBER: US/09/764,877
: CURRENT FILING DATE: 2001-01-17
: Prior application data removed
: NUMBER OF SEQ ID NOS: 4031
SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1727
: LENGTH: 57
: TYPE: PRT
: ORGANISM: Homo sapiens
FEATURE:
: NAME/KEY: SITE
: LOCATION: (1)
: OTHER INFORMATION: Xaa equals any of the naturally occur
: NAME/KEY: SITE
: LOCATION: (7)
: OTHER INFORMATION: Xaa equals any of the naturally occur
: NAME/KEY: SITE
: LOCATION: (36)
: OTHER INFORMATION: Xaa equals any of the naturally occur
US-09-764-877-1727

Query Match          1.8% Score 6; DB 10; Length 5
Best Local Similarity 100.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 QLGILG 16
DB 8 QLGILG 13

RESULT 33
US-09-781-988-78
: Sequence 78, Application US/09781988
: Patent No. US20020150881A1
GENERAL INFORMATION:

```

```

APPLICANT: Laddner, Robert Charles
            Guterman, Sonia Kosow
            Poverly, Rocco Lindsay
            Markland, William
            Ley, Arthur Charles
            Kent, Rachel Baribault
TITLE OF INVENTION: Directed Evolution of No. US20020150841A1C1
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESS: Broadway and Nelmark
STREET: 419 Seventh Street, N.W.
Suite 300
CITY: Washington,
STATE: DC
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 4.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/781,988
FILING DATE: 14-Feb-2001
CLASSIFICATION: ~Unknown~
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/664,989
FILING DATE: ~Unknown~
APPLICATION NUMBER: 07/487,063
FILING DATE: 02-MAR-1990
APPLICATION NUMBER: 07/240,160
FILING DATE: 02-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Casper, Iveta P
REGISTRATION NUMBER: 28005
REFERENCE/PRIORITY NUMBER: 1A1NFP 7
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 628 5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 78:
SEQUENCE CHARACTERISTICS:
LENGTH: 58 amino acids
TYPE: amino acid
MOLECULE TYPE: protein
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 78:
US-09-781-988-78

Query Match          1.8% Score 6; FR 10; Length 58;
Best Local Similarity 100.0%; Pred No. 81,
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 QCLGHC 21
DB 37 QCLGHC 42

RESULT 34
US-09-867-550-1392
Sequence 1392, Application US/9987550
Patent No. US20020082206A1
GENERAL INFORMATION:
APPLICANT: Leach, Martin D.
APPLICANT: Mehraban, Fuad,
APPLICANT: Gouliay, Pamela
APPLICANT: Law, Debbie
APPLICANT: Topper, James
TITLE OF INVENTION: No. US20020082206A1C1 Polymers of peptides from Alphetocids cells and
FILE REFERENCE: 21402-013 (Gura-313)
CURRENT APPLICATION NUMBER: US/09/867,550
CURRENT FILING DATE: 2001-09-20

```

```

PRIORITY APPLICATION NUMBER: USSN 60/208,427
PRIORITY FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 2125
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1392
LENGTH: 90
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)
OTHER INFORMATION: N. Mergin Xia may be any one of Ala or Pro as set or Thr
US-09-867-550-1392

Query Match          1.8% Score 6; DB 10; Length 90;
Best Local Similarity 100.0%; Pred No. 1,2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 271 SPCTGC 276
DB 81 SPCTGC 86

RESULT 35
US-10-004-381-17
Sequence 17, Application US/10004381
Patent No. US20020155578A1
GENERAL INFORMATION:
APPLICANT: SESTAK, JACK W.
APPLICANT: WILSON, DAVID S.
APPLICANT: KEEPE, ANTHONY D.
TITLE OF INVENTION: STEPTAVIUMIN-BINDING PEPTIDES AND USES
FILE REFERENCE: 00786/388002
CURRENT APPLICATION NUMBER: US/10/004,381
CURRENT FILING DATE: 2001-10-31
PRIORITY APPLICATION NUMBER: US 79/244,541
PRIORITY FILING DATE: 2000-10-31
NUMBER OF SEQ ID NOS: 41
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 17
LENGTH: 102
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: selected peptide
US-10-004-381-17

Query Match          1.8% Score 6; DB 9; Length 102;
Best Local Similarity 100.0%; Pred No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 QCLGHC 16
DB 80 QCLGHC 85

RESULT 36
US-09-764-853-786
Sequence 786, Application US/09/764854
Patent No. US20020090672A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, proteins, and Antibodies
FILE REFERENCE: P1206
CURRENT APPLICATION NUMBER: US/09/764,853
CURRENT FILING DATE: 2001-01-17
PRIORITY APPLICATION data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 939
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 786
LENGTH: 111
TYPE: PRT

```

```

: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (53)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (101)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-853-786

```

```

Query Match          1.88: Score 6; DB 10; Length 11;
Best Local Similarity 100.0%; Pred. No. 1 4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 32 LLAGVL 37
    |||||
DB 69 LLAGVL 74

```

```

RESULT 37
US-09-864-761-45806
: Sequence 45806, Application US09098476.1
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chu, W-sheng
: TITLE OF INVENTION: HUMAN GENOME-DEPENDED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: FILE REFERENCE: Aeomica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: PRIOR FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/642,396
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: CH 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO: 45806
: LENGTH: 117

```

```

: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC021880.2
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.49
: OTHER INFORMATION: SWISSPROT HIT: P04652, EVALU8 9.00e-
: OTHER INFORMATION: EST_HUMAN HIT: AW81235.1, EVALU8 3.0
US-09-864-761-45806

```

```

Query Match          1.88: Score 6; DB 10; Length 11;
Best Local Similarity 100.0%; Pred. No. 1 5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 180 LIDLKT 185
    |||||
DB 20 LIDLKT 25

```

```

RESULT 38
US-09-764-870-302
: Sequence 302, Application US09764470
: Patent No. US20020042386A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PT214
: CURRENT APPLICATION NUMBER: US/09/764,870
: PRIOR FILING DATE: 2001-01-17
: PRIOR APPLICATION data removed Consult PALM or file with
: NUMBER OF SEQ ID NOS: 646
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO: 302
: LENGTH: 119
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-764-870-302

```

```

Query Match          1.88: Score 6; DB 10; Length 11;
Best Local Similarity 100.0%; Pred. No. 1 5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 266 QWVDS 271
    |||||
DB 91 QWVDS 96

```

```

RESULT 39
US-09-815-242-12092
: Sequence 12092, Application US09815242
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert L.
: TITLE OF INVENTION: Identification of Essential Genes in
: FILE REFERENCE: ELITRA.011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: PRIOR FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: US/09/815,242
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-25
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-24
: PRIOR APPLICATION NUMBER: 60/254,625

```

```

: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO: 12092
: LENGTH: 131
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-815-242-12092

Query Match
Best Local Similarity: 100.0%; Score 6; DB 10; Length 131;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SKEPRV 9
DB 46 SKEPRV 51

RESULT 40
US-09-216-393-66
: Sequence 66, Application US/09216393
: Patent No. US2001001447A1
: GENERAL INFORMATION:
: APPLICANT: Milhausen, Michael James
: TITLE OF INVENTION: TOXOPLASMA GONDI PROTEINS, NUCLEIC ACID MOLECULES, AND
: FILE REFERENCE: US/09/216,393
: CURRENT APPLICATION NUMBER: US/09/216,393
: PRIOR FILING DATE: 1998-12-18
: EARLIER APPLICATION NUMBER: 08/994,825
: PRIOR FILING DATE: 1997-12-19
: NUMBER OF SEQ ID NOS: 364
: SOFTWARE: Patent In Ver. 2.0
: SEQ ID NO: 66
: LENGTH: 138
: TYPE: PRT
: ORGANISM: Toxoplasma gondii
US-09-216-393-66

Query Match
Best Local Similarity: 100.0%; Score 6; DB 10; Length 138;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 SLSOE 54
DB 45 SLSOE 50

RESULT 41
US-09-815-242-10702
: Sequence 10702, Application US/09815242
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlson, Karl L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: TITLE OF INVENTION: Identification of Essential Genes in
: FILE REFERENCE: FLITRA.011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: PRIOR FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,074
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848

```

```

: PRIOR FILING DATE: 2000-05-24
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO: 10702
: LENGTH: 139
: TYPE: PRT
: ORGANISM: Enterococcus faecalis
US-09-815-242-10702

Query Match
Best Local Similarity: 100.0%; Score 6; DB 10; Length 139;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 QELTOL 91
DB 48 QELTOL 53

RESULT 42
US-09-764-868-1057
: Sequence 1057, Application US/09764868
: Patent No. US2002016871A1
: GENERAL INFORMATION:
: APPLICANT: Hesse et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antilodies
: FILE REFERENCE: PTE22
: CURRENT APPLICATION NUMBER: US/09/764,868
: PRIOR FILING DATE: 2001-01-17
: PRIOR APPLICATION data removed - refer to PALM or file wrapper
: NUMBER OF SEQ ID NOS: 1510
: SOFTWARE: Patent In Ver. 2.0
: SEQ ID NO: 1057
: LENGTH: 153
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-764-868-1057

Query Match
Best Local Similarity: 100.0%; Score 6; DB 9; Length 153;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 LAGVAV 48
DB 120 LAGVAV 125

RESULT 43
US-09-825-301-27
: Sequence 27, Application US/09825301
: Patent No. US20020009738A1
: GENERAL INFORMATION:
: APPLICANT: Houghton, Raymond L.
: APPLICANT: Dillon, David C.
: APPLICANT: Molesh, David A.
: APPLICANT: Xu, Jiaqun
: APPLICANT: Zehner, Barbara
: APPLICANT: Persing, David H.
: TITLE OF INVENTION: METHODS, COMPOSITIONS AND KITS FOR THE DETECTION
: FILE REFERENCE: 210121.513
: CURRENT APPLICATION NUMBER: US/09/825,301
: PRIOR FILING DATE: 2001-04-02
: NUMBER OF SEQ ID NOS: 77

```



```

: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO: 27
: LENGTH: 154
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (148)
: OTHER INFORMATION: Xaa - Any Amino Acid
US-09-825-301-27

```

```

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 154;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 157 LTRKA 162
DB 33 LTRKA 38

```

RESULT 44

```

US-09-604-287A-471
: Sequence 471, Application US/09604287A
: Patent No. US20020064872A1
: GENERAL INFORMATION:
: APPLICANT: Jiang, Yugu
: APPLICANT: Dillon, David C.
: APPLICANT: Mitcham, Jennifer L.
: APPLICANT: Xu, Jiangchun
: APPLICANT: Harlocker, Susan L.
: APPLICANT: Hepler, William T.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
: FILE REFERENCE: 210121.470C7
: CURRENT APPLICATION NUMBER: US/09/604.287A
: CURRENT FILING DATE: 2000-06-22
: NUMBER OF SEQ ID NOS: 489
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO: 471
: LENGTH: 154
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (148)
: OTHER INFORMATION: Xaa - Any Amino Acid
US-09-604-287A-471

```

```

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 154;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 157 LTRKA 162
DB 33 LTRKA 38

```

RESULT 45

```

US-10-007-805-471
: Sequence 471, Application US/10007805
: Patent No. US20020150581A1
: GENERAL INFORMATION:
: APPLICANT: Jiang, Yugu
: APPLICANT: Dillon, David C.
: APPLICANT: Mitcham, Jennifer L.
: APPLICANT: Xu, Jiangchun
: APPLICANT: Harlocker, Susan L.
: APPLICANT: Hepler, William T.
: APPLICANT: Henderson, Robert A.
: APPLICANT: Fanger, Gary R.
: APPLICANT: Vedvick, Thomas S.
: APPLICANT: McNeill, Patricia D.
: APPLICANT: Durham, Margarita

```

```

: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: FILE REFERENCE: 210121.470C10
: CURRENT APPLICATION NUMBER: US/10/007.805
: CURRENT FILING DATE: 2001-12-07
: NUMBER OF SEQ ID NOS: 593
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO: 471
: LENGTH: 154
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: VARIANT
: LOCATION: 148
: OTHER INFORMATION: Xaa - Any Amino Acid
US-10-007-805-471

```

```

Query Match
Best Local Similarity 100.0%; Score 6; DB 12; Length 154;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 157 LTRKA 162
DB 33 LTRKA 38

```

```

Search completed: December 7, 2002, 11:18:48
Job time : 18 secs

```

...


```
/organism="Homo sapiens"  
/db_xref="taxon:9606"  
/clone="IMAGF:5166966"  
/clone_id="NH_MGC_119"  
/clone_type="modul1a"  
/lab_host="DH10B"
```

note: Organ: brain; Vector: JMW-SH0816; Site: 1; Not1 Site: 2; EcoPvI (test/repeat): PNA source: normal meth1a from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9–3 kb. Library is normalized and enriched for full-length clones and was transfected by Ca²⁺ transfer (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH-MGC library.

BASE COUNT	204 a	193 c	235 g	137 t
ORIGIN				
Query Match	29.98%	Score 461	DB 133	Length 764
Best Local Similarity	100.00%	Fred No. 3	1e-221	
Matches 461; Conservative	0;	Mismatches	0;	Indels 0;
				Gaps 0;

[illegible]

RESULT 2	LOCUS	DEFINITION	ACCESSION VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
AL544658	AL544658	792 bp, mRNA, linear	AL544658	prime, mRNA sequence.	human.	<i>Homo sapiens</i>	Eukaryota; Metazoa; Chordata; Vertebrata; Euarchontomi; Mammalia; Euliefia; Primates; Catarrhini; Hominoidea; Homo.	Li, W. B., Gruber, C., Tesse, J. and Polayes, D.	Full-length cDNA libraries and normalization	unpublished (2001)	contact: Genoscope, Centre National de Séquençage

REF 191 91006 EVRY Index - France
 Email - semp@francevoies.org - Chat - Web - www.francevoies.org/vis1.fr
 Location/Qualifiers
 1. 792

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="pSub102YH14"
/clone_lib="UT_NFL06_PL2"
/issue_type="Placenta"
/issue_tissue="Placenta"

```

Zhong et al., p-STAT1 gene expression profile
Zhang et al., p-STAT1 gene expression profile

Five prime end enriched, double stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pGEMSPORT 6 vector. Library was normalized. Library was constructed by life technologies; contact : Feng Liang life technologies a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 619 8371 Email : fliang@life.com url : http://full-length.invitrogen.com"

Base 3500T
Origin

208 a 202 c 254 g 146 t 2 others

[illegible]

RESULT 3	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
AV645755	796 bp	AV645755 GAT Homo sapiens cDNA clone G1CAED05 3', mRNA sequence.	AV645755	AV645755.1	GI:9866769	EST.	human.		
							Homo sapiens		
							Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
							Mammalia; Eutheria; primates; catarrhini; hominidae; Homo.		
							(Buses 1 to 706)		
							Xu, X., Huang, D., Xu, Z., Qian, B., Zhu, Z., Yan, G., Cai, T., Zhang, X.,		
							Xiao, H., Guo, T., Liu, F., Huang, Q., Cheng, Z., Li, N., Du, J., Hu, W.,		

TITLES

MEIDIANI

COMMENT

Shen, K., Gu, J., Chen, Z. and Han, Z.
Hu, G., Gu, J., Chen, Z. and Han, Z.
Insight into hepatocellular carcinogenesis at transcriptional level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
21625106

Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shouling Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex. 45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn

This clone is available at CHGC in Shanghai.

FEATURES
SOURCE

BASE COUNT	168 a	198 c	176 g	162 t	2 others
ORIGIN					

Query Match	25.0%	Score 385;	DB 10;	Length 706;
Best Local Similarity	100.0%	Pred. No. 6.8e-183;		
Matches 385:	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

RESULT 4	BE562903	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
BE562903	601336339P1	R28 bp	mRNA	11069	EST	15 AUG-2000		
BE562903	601336339P1	NH_MGC_44	Homo sapiens cDNA	clone IMAGE:360319.57,	mRNA Sequence.			
BE562903.1	GI:9806623	EST	human.					
Human.								
Homo sapiens								
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (pages 1 to 82)	NH-M&P	http://mc.manuscriptcentral.com/nhm	National Institutes of Health, Mammalian Gene Collection (1999)	Contact: Robert Strausberg, Ph.D.

FEATURES

Sources

BASE COUNT	232 3	208 5	244 9	144 t
ORIGIN				

Query Match	24.58;	Score 378;	EB 10;	Length 1
Best Local Similarity	100.08;	Prod. No. 2,36-179;		
Matches 378;	Conservative 0;	Mismatches 0;	Indel.	

QY	397	AGGTTTAAAGGTCCTCCAGCTCTCTTACGACAGCAATCTCCAGCAATAGC	100	100
Db	117	AGGTCCAAAGTCCCGACGCTGCTTAAGTATGAGAAATATCCGAGCAAGAC	100	100
QY	457	GAACCTGACCCAGCTTAAAGCTGCACTGAGTCAAGTCTCCAGAGAAATCCCA	100	100
Db	177	GAACCTGACCCAGCTTAAAGCTGCACTGAGTCAAGTCTCCAGAGAAATCCCA	100	100
QY	517	GATCTACCAAGGAGCTGACCTAGTCTGAAGGCTGATGAGGTTGGTAT	100	100
Db	237	GATCTACCAAGGAGCTGACCTAGTCTGAAGGCTGATGAGGTTGGTAT	100	100
QY	577	GCTGCAGCATGATTCACCAAGTATGCTGAGCTGCAAGCTGCAGTGGCTG	100	100
Db	297	GCTGCAGCATGATTCACCAAGTATGCTGAGCTGCAAGCTGCAGTGGCTG	100	100
QY	637	GAATTCGAAGCTCCAGAGAGATTAACATAGAGTGCACCAAGCTGAAGCTG	100	100
Db	357	GAATTCGAAGCTCCAGAGAGATTAACATAGAGTGCACCAAGCTGAAGCTG	100	100
QY	697	GTTGCCAGAGAAATCCCAAGTCTGAGTACATCTAGTACGACAGCTGATCTGG	100	100
Db	417	GTTGCCAGAGAAATCCCAAGTCTGAGTACATCTAGTACGACAGCTGATCTGG	100	100
QY	757	AGTGGTGAATGTCGAGA	774	
Db	477	AGTGGTGAATGTCGAGA	494	

RESULT 5	LOCUS	DEFINITION	ACCESSION
B1827055	B1827055	786 bp, mRNA, 1116 bp, cDNA clone	60307572661 NIH_Myc_119 Homo sapiens cDNA clone
B1827055	B1827055	mRNA sequence.	B1827055

[illegible][illegible]

```

QY 1010 ACTAACGCTTCTCTGATGGACATTTCATATTAATTAAGAAAGAAATGGAATG 1069
DB 246 AGTAACCCCTCTCTGATGGATGATATTTTCAGACCTAAATAGTAAGTACAGTGCATG 305
QY 1070 GTGGAGGCTACCTCTCTGATGGATGATATTTTCAGACCTAAATAGTAAGTACAGTGCATG 1129
DB 306 GTGGAGGCTACCTCTCTGATGGATGATATTTTCAGACCTAAATAGTAAGTACAGTGCATG 365
QY 1140 AATAGCGGGAATGAGACTCTG 1151
DB 366 AATAGCGGGAATGAGACTCTG 367

RESULT 7
LOCUS A0140165 792 bp mRNA linear EST 05-AUG-2002
DEFINITION A0140165 PLACE2 Homo sapiens cDNA clone PLACE2000062 5', mRNA
ACCESSION A0140165
VERSION A0140165.1 GI:11001686
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
REFERENCE 1 (bases 1 to 792)
AUTHORS Ota,T., Suzuki,Y., Saito,K., Ishii,S., Yamamoto,J., Sugiyama,T.,
Nishikawa,T., Nakamura,Y., Sugano,S., Masuho,Y., and Isegai,T.
HRI human cDNA project (Ota,T., Suzuki,Y., Saito,K., Ishii,S.,
Yamamoto,J., Sugiyama,T., Nishikawa,T., Nakamura,Y., Sugano,S.,
Masuho,Y., Isegai,T.)
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isegai
Genomics Laboratory
Helix Research Institute
1512-3 Yana, Kisarazu, Chiba 299-8812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: gromic@helix.co.jp
HRI human cDNA project: 5'- & 3'-end one pass sequence ng: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
Source 1..792
Location/Qualifiers
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone="PLACE2000062"
  /clone_id="PLACE2"
  /issue_type="placenta"
  /note="Vector: pME18SFL3"

BASE COUNT 214 a 191 c 234 g 149 t 4 others

Query Match 17.88: Score 275, DB 9, Length 792,
Best Local Similarity 94.58: Pred. No. 2,1e-127,
Matches 375: Conservative 0; Mismatches 2; Indels 3; Gaps 0;

```

```

QY 638 AATTCGAAGCTGAGGAGATCTATCAGAGAGTGAACCGGCTGAAGGCTG 1069
DB 293 AATTCGAAGCTGAGGAGATCTATCAGAGAGTGAACCGGCTGAAGGCTG 1069
QY 698 TTGCGAGAGAAATCCAAAGTGTATATATCTAGCGAGAGCTGACCGAG 1129
DB 353 TTGCGAGAGAAATCCAAAGTGTATATATCTAGCGAGAGCTGACCGAG 1129
QY 758 GTGGGTGAGTGGCAGA 774
DB 413 GTGGGTGAGTGGCAGA 429

RESULT 8
LOCUS AAT79963 326 bp mRNA linear EST 1998
DEFINITION 238608.81 Soares_fetal_liver_spleen_INFLS.St.H.
ACCESSION AAT79963
VERSION AAT79963.1 GI:2844294
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae
REFERENCE 1 (bases 1 to 326)
AUTHORS Hillier,L., Alton,M., Howles,L., Dubouque,T., Gels
Krizman,D., Kuehna,T., Lacy,M., Le,N., Lennon,G.,
J., Moutre,B., Schellenberg,K., Steptoe,M., Tan,
White,Y., Wylie,T., Waterson,K., and Wilson,R.
WashU-NCI human EST project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson R
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNC
IMAGE Consortium (info_image_lnc@wustl.edu) for full
Seq primer: -40m13 fwd: RT from Amersham
High quality sequence stop: 246.
FEATURES
Source 1..326
Location/Qualifiers
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /db_xref="IMAGE:461966"
  /clone="IMAGE:461966"
  /clone_id="Soares_fetal_liver_spleen_13"
  /sex="male"
  /des_state="20 week post conception fo"
  /lab_host="Dh10B (ampicillin resistant)"
  /note="Organ: Liver and Spleen, Vector:
  with a modified polylinker site. 1: Part
  This is a subtracted version of the or"
  liver spleen INFLS library. 1st stran"
  with a Pac 1 - oligo(dT) primer (5'
  AACTCGAAGAAATTAATTAAGATCTTTT1111111
  double-stranded cDNA was ligated to Eco
  (Pharmacia), digested with Pac 1 and c
  and Eco RI sites of the modified f177s
  went through one round of normalization,
  constructed by Bento Soares and M. Patric"

BASE COUNT 93 a 64 c 97 g 82 t

Query Match 15.68: Score 241, DB 9, Length
Best Local Similarity 100.0%, Pred. No. 3,3e-110,
Matches 241: Conservative 0; Mismatches 0; Indels 4;

```


Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p7713-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dfr)18 tail. The sequence tag for this library is

ACACTTCGAC.
TAG_L1B-01-H-E11
TAG_R1SSD-chondrosarcoma
TAG_SEQ-ACACTTCGAC"

HASE COUNT 202 a 139 c 173 g 142 t

ORIGIN

Query Match 14.8% Score 228; DB 14; Length 656;
Best Local Similarity 100.0%; Pred.No. 1,1e-103;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1302 CTACACCGCGGACACCTTTGATGGGAGATTGCTGTTGAAGAGACTATCTTGAGG 1361
|||||
DB 244 CTACACCGCGGACACCTTTGATGGGAGATTGCTGTTGAAGAGACTATCTTGAGG 185
QY 1362 CTGAGAGCAGCAGGGAATTAATCCCGAGTCAATGAGTCAAGAGAGAGAGAG 1421
|||||
DB 184 CTGAGAGCAGCAGGGAATTAATCCCGAGTCAATGAGTCAAGAGAGAGAGAGAG 125
QY 1422 GAGTCATACACCTCTCCACCGCCGACATCTTCTTGTCTAATAGTCTTCTT 1481
|||||
DB 124 GAGTCATACACCTCTCCACCGCCGACATCTTCTTGTCTAATAGTCTTCTT 65
QY 1482 GAGTCATACACCTCTCCACCGCCGACATCTTCTTGTCTAATAGTCTTCTT 1529
|||||
DB 64 GAGTCATACACCTCTCCACCGCCGACATCTTCTTGTCTAATAGTCTTCTT 17

RESULT 11

LOCUS

H91270/c

DEFINITION

yu99e03.s1 Soares fetal liver spleen INFILS Homo sapiens cDNA clone

ACCESSION

H91270

VERSION

H91270.1

KEYWORDS

EST.

SOURCE

ORGANISM

human.

REFERENCE

AUTHORS

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hulman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rikhi, L., Rohlfing, T., Soares, M., Tan, F., Trevisan, E., Waterston, R., Williamson, A., Wohlmann, P., and Wilson, R.

TITLE

JOURNAL

COMMENT

Unpublished (1995)

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Insert Size: 732

High quality sequence stops: 193

Source: IMAGE Consortium; LINT

This clone is available royalty free through LINT. Contact: The

IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 732 Std Error: 0.00

Seq primer: Promega -21ml3

High quality sequence stops: 193

Location/Qualifiers

1..289

/organism="Homo sapiens"

/db_xref="cds:3790505"

/db_xref="taxon:9606"

/clone="IMAGE:241372"

/clone="lib-Soares fetal liver spleen 18"

/sex="male"

/dev_stage="29 week post conception test"

/lab_host="pB10H (ampicillin resistant)"

/note="Organ: Liver and Spleen; Vector: p"

with a modified polylinker. Site 1: Pac

1st strand cDNA was primed with a Pac 1

15' ACTGGAATATTAATTAAGATCTTTTCTT

double-stranded cDNA was ligated to Eco

(Pharmacia), digested with Pac 1 and cDNA

and Eco RI sites of the modified p7713

went through one round of normalization

constructed by Benito Soares and M. Farina

BASE COUNT 82 a 58 c 84 g 63 t

ORIGIN

Query Match 14.6% Score 225; DB 14; Length
Best Local Similarity 100.0%; Pred.No. 3.9e-102;
Matches 225; Conservative 0; Mismatches 0; Indels 0;

QY 1306 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1421
|||||
DB 225 ACCCGTGACCGCTTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1421
QY 1366 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1421
|||||
DB 165 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1421
QY 1426 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1421
|||||
DB 105 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1421
QY 1486 GTTCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1530
|||||
DB 45 GTTCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1530

RESULT 12

LOCUS

AA448002/c

DEFINITION

2962912.s1 Soares fetal liver spleen INFILS Homo sapiens cDNA clone

ACCESSION

AA448002

VERSION

AA448002.1

KEYWORDS

EST.

SOURCE

ORGANISM

human.

REFERENCE

AUTHORS

Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Parsons, J., Rikhi, L., Rohlfing, T., Soares, M., Tan, F., Trevisan, E., Waterston, R., Williamson, A., Wohlmann, P., and Wilson, R.

TITLE

JOURNAL

COMMENT

Unpublished (1997)

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty free through LINT.

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -41ml3 fwd. RT from Amersham

High quality sequence stops: 416.

Location/Qualifiers

1..506

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:742758"

/clone="lib-Soares fetal liver 18"

/sex="male"

/clone="IMAGE:240697"
 /clone.lib="Scores fetal liver spleen INFLS"
 /sex="male"
 /dev_stage="20 week post conception fetus"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Organ: Liver and Spleen; Vector: p7133 (Pharmacia)
 with a modified polylinker; Site:1 Pac I site;2 Eco RI
 1st strand cDNA was primed with a Pac I - oligo(dT) primer
 15' AACGAGCAATTAATTAAGACCTTTTCTTTTCTTTTCTTT
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Pac I and cloned into the Pac I
 and Eco RI sites of the modified p7133 vector. Library
 went through one round of normalization. Library
 constructed by Bento Soares and M. Patricia Bonaldo."

BASE COUNT 107 a 85 c 126 g 90 t 2 others
 ORIGIN

Query Match 10.6%; Score 164; DB 14; Length 410;
 Best Local Similarity 100.0%; Freq. No. 2,26,71.
 Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

Y 169 CACACGCAATTAATTTCTTCTATTAACCAATGATCTAAATGAGATGATGCTTA 1428
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 164 CACAGGCAATTAATTTCTTCTATTAACCAATGATCTAAATGAGATGATGCTTA 105
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Y 1429 TACAACCTTCGACCCGACATCTTTTGTGTCTATACATCTCTTCCATTTGGCTGTT 1468
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 104 TACAACCTTCGACCCGACATCTTTTGTGTCTATACATCTCTTCCATTTGGCTGTT 45
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Y 1489 TCTGAGTTGAGCCTTTATTAATGAAGTGGTAATTTTACTG 1532
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 44 TCTGAGTTGAGCCTTTATTAATGAAGTGGTAATTTTACTG 1

RESULT 15
 AL570816/c
 LOCUS
 DEFINITION L1L_NFL006_P12 Homo sapiens cDNA clone CS001022YH14.3
 AL570816 prime, mRNA sequence.
 AL570816
 VERSION AL570816.1 GI:12927495
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 11,118 bp. Gruber, C., Jesse, J. and Polayars, D.
 Full-length cDNA libraries and normalization
 unpublished (2001)

COMMENT
 Genoscope: Centre National de Sequencage
 R# 191 91006 EVRY cedex - France
 Email: sege@genoscope.cns.fr; Web: www.genoscope.cns.fr.
 Location/Qualifiers
 1..833
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS001022YH14"
 /clone.lib="L1L_NFL006_P12"
 /tissue_type="Placenta"
 /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6
 vector. Library was normalized. Library was constructed by
 life Technologies. Contact: Peng Liang Life Technologies,
 a division of Invitrogen 4800 Medical Center Drive
 Rockville, Maryland 20850, USA Fax: (1) 301 610 8371
 Email: tliang@life-tech.com URL:
 http://fulllength.invitrogen.com"

FEATURES
 SOURCE

CDS: 233 a 193 c 219 g 184 t 4 others

[illegible]

BASE: COUNTY
ORIGIN

1153	1220	1169	1221	9 Oct. 1910
------	------	------	------	-------------

SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota, Metazoa, Chordata, Mammalia, Vertebrata; Euteleostomi; Mammalia, Euteria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 499)
AUTHORS	Hillier,L., Allen,M., Nowles,L., Dubouque,T., Geisler,G., Jost,S., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,R., White,Y., Wyllie, T., Waterston,K. and Wilson,R.
TITLE	WashU-Merck EST project 1997
JOURNAL	Unpublished (1997)
COMMENT	Contact: Willson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.wustl.edu This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Seq primer: -28m3 rev2 ET from Amersham High quality sequence stop: 495. Location/Qualifiers
FEATURES	1..499
SOURCE	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:782332" /sex="male" /lab_host="DH10H" /note="Vector: pUT730-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories , Inc., and primed with a Not I - oligo(dT) primer [5' TGTTTCACATCTGAAGTAGGAGCGGCCGCACATTGTTTTTTTTTT 3'] Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pUT73 vector. Library went through one round of normalization to Cc55, and was constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT	94 A 135 C 112 G 158 T
ORIGIN	
Query Match:	6.7% Score 104; DB 9; Length 499;
Best Local Similarity	100.0%; Pred. NO. 4; 1e-41;
Matches 104; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Gy 1198	CATTACTGGCATCTGCAAAAGCCCGCAGCCCTGCTAGAGACAATAGTGTTCCTG 1257
Dd 1	CAATTACTGATGTGCAAAAAGCCCGCAGCCCTGCTAGAGACAATAGTGTTCCTG 60
Gy 1258	CTAGCCCTACGCTCCATCTGTGTAATAGAGAACATTCACCCACTT 1301
Dd 61	CTAGCCCTACGCTCCATCTGTGTAATAGAGAACATTCACCCACTT 104
RESULT 24	
LOCUS	H009487 393 bp mRNA linear EST 19-JUN-1995
DEFINITION	yj28604.s1 Soares placenta NiZHP Human sapiens cDNA clone
ACCESSION	IMAGE:150055.3, mRNA sequence.
VERSION	H009487
KEYWORDS	H009487.1 GI:863920
SOURCE	EST.
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota, Metazoa, Chordata, Mammalia, Vertebrata; Euteleostomi, Mammalia, Euteria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 393) Hillier,L., Clark,N., Dubouque,T., Lennon,K., Hawkins,M., Holman, M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Roling,T., Soares,M., Tan,F., Trevisakis,Z., Waterston, R., Williamson,A., Wohlmann,P. and Wilson,R.
TITLE	The WashU-Merck EST Project

JOURNAL	Unpublished (1995)
COMMENT	Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway Box 8601, St. Louis, MO Tel: 314 286 1800 Fax: 314 286 1810 Email: estewatson.wustl.edu Insert Size: 662
FEATURES	High quality sequence steps: 297 Source: IMAGE Consortium, LLN. This clone is available royalty-free through LNC IMAGE Consortium (info:image.lln.gov) for further insert length: 662 Std Error: 0.00 Seq primer: Promega -21mt+ High quality sequence step: 297.
SOURCE	location/quality 1. 493 /organism "Homo sapiens" /db_xref-"dbx:661994" /db_xref-"taxon:9606" /clone-"IMAGE:150055" /clone_lib-"Soares placenta NB2HP" /sex-"female" /dev_stage "Placenta obtained at birth /lab host "PHIB (ampicillin resistant) /note-"primate placenta; Vector: pTZ19 modified polylinker; Site 1: Not 1; Site strand cDNA was primed with a Not 1 site AACTGCAATTCATCTCCCGCAGACAAATTTTCTTT double-stranded cDNA was ligated to Eco (Pharmacia), digested with Not 1 and Xba I and Eco RI sites of the modified pTZ19 went through one round of normalization constructed by Renato Soares and M.Patino
BASE COUNT	111 a 84 c 104 g 84 t 11 other
ORIGIN	
Query Match	6.08; Score 94; DB 14; Length
Best Local Similarity	99.3%; Pred. No. 1; %e=35;
Matches 143; Conservative	0; Mismatches 1; Indel:
Oy 1384	TCCCCTAGTCATTCAATGGTAATGCATAGAGGAGGCTCATACAA Db 150 TCCCCAATCAAAACAATGAATGTAAACAGAGGAGGCTCATACAA Oy 1444 CCCCACATCTTCTTGCTCTATAATATCTTGCATTTGGCTGTTTCTGA Db 90 CCCCACATCTTCTTGCTCTATAATATCTTGCATTTGGCTGTTTCTGA Oy 1504 TTATATAAGTGTAAATCTCT 1527 Db 30 TTATATAAGTGTAAATCTCT 7
RESULT 25	
AAd002266/c	
LOCUS	AAd002266 428 bp mRNA linear
DEFINITION	ZNR82003.S1 Soares total liver spleen INFIS_S1 full
ACCESSION	U082003.1
VERSION	AA002266
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota: Chordata: Granulata: Vertebrata: Mammalia: Eutheria: Primates: Catarrhini: Homini: 1 (bases 1 to 428) Hallier,L., Lennon,G., Bekke,M., Bonaldo,M.F., Chissoe,S., Dietrich,N., DuBouye,J., Favello,A., M., Holtman,M., Kuraba,T., Lacy,M., Le,M., Le,B., R., Morris,M., Parsons,J., Prade,C., Riklin,I., Schellenberg,K., Soares,M.B., Tan,F., Thierry-M, Underwood,K., Wohlmann,P., Watson,R., Willet,

Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer genome project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/getitem2.pl?1-55-11-2707001-0504-1)
 Seq primer: puc 18 forward
 High quality sequence start: 78
 High quality sequence stop: 254
 Location/Qualifiers
 1..254
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="TNO025"
 /dev_stage="Adult"
 /note="Organ: testis-normal; Vector: puc18; Site 1: Small; Site 2: Small; A mini-library was made by cloning products derived from OKESHS PCR (0.5 letters patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT
 ORIGIN
 59 a 67 c 56 g 72 t

Query Match 2.28; Score 34; EB 12; Length 254;
 Best Local Similarity 100.0%; Pred. No. 9; E=0.06;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1228 CTGCTTCAGACGACATAGTGTTCCTGCTAG 1251
 DB 94 CTGCTTCAGACGACATAGTGTTCCTGCTAG 127

RESULT 28
 AA918594 405 bp mRNA linear EST 10-JUN-1998
 LOCUS 0153c12.s1 Soares_NFL_I_GRC_S1 Homo sapiens cDNA clone
 DEFINITION IMAGE:1527190.3, mRNA sequence.
 AA918594
 VERSION AA918594.1 GI:3058484
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 405)
 REFERENCE NCI-CCAF http://www.ncbi.nlm.nih.gov/ncicgap.
 Tumor Gene Index
 Tumor Gene Index
 Unpublished (1997)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 This clone is available royalty-free through LIN: contact the
 IMAGE Consortium (info@image.jiml.gov) for further information.
 Insert Length: 513 Std Error: 0.00
 Seq primer: 40ml3 fwd. ET from Amersham
 High quality sequence stop: 338.
 Location/Qualifiers
 1..405
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="IMAG:1527190"
 /clone_lib="Soares_NFL_I_GRC_S1"
 /lab_host="DH10B"
 /note="Organ: pooled; Vector: p7770-pac (Pharmacia) with
 a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
 Equal amounts of plasmid DNA from three normal
 libraries (fetal lung NBH19W, testis NHT, and B-cell
 NCI-CCAF) were mixed, and ss circles were made in
 vitro. Following HAP purification, this DNA was used as
 tracer in a subtractive hybridization reaction. The direct

was PCR-amplified cDNAs from pools of
 from the same 3 libraries. The pools con-
 tain M.A.G.E. clones 297480, 620887, 682642,
 726408-728711, and 729996-731399. Submitted
 Soares and M. Fatima Honaido.
 BASE COUNT 100 a 88 c 89 g 128 t
 ORIGIN

Query Match 2.28; Score 34; EB 3; Length 4
 Best Local Similarity 100.0%; Pred. No. E=0.06;
 Matches 34; Conservative 0; Mismatches 0; Indels 0;

OY 1479 TTGCTCTTCTCAGTGTGACCTGATTAATAA 1512
 DB 289 TTGCTCTTCTCAGTGTGACCTGATTAATAA 22

RESULT 29
 AA1217955 453 bp mRNA linear EST 1998
 LOCUS q152c6.x1 Soares_testis_NHT Homo sapiens cDNA
 DEFINITION IMAGE:1753642.3, mRNA sequence.
 AA1217955
 VERSION AA1217955.1 GI:3797770
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
 1 (bases 1 to 453)
 REFERENCE NCI-CCAF http://www.ncbi.nlm.nih.gov/ncicgap.
 Tumor Gene Index
 Tumor Gene Index
 Unpublished (1997)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 cDNA library preparation: M. Bento Soares, Ph.D.
 , Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome
 Clone distribution: NCI-CCAF clone distribution
 found through the M.A.G.E. Consortium/LIN: at
 www.bio.liml.gov/ftp/image/image.html
 Insert Length: 548 Std Error: 0.00
 Seq primer: 40UP from Gibco.
 Location/Qualifiers
 1..453
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="IMAG:1753642"
 /clone_lib="Soares_testis_NHT"
 /seq_host="DH10B"
 /lab_host="DH10B"
 /note="Vector: p7770-pac (Pharmacia) with
 a modified polylinker; Site 1: Not 1; Site 2: Eco
 RI; Equal amounts of plasmid DNA from three normal
 libraries (fetal lung NBH19W, testis NHT, and B-cell
 NCI-CCAF) were mixed, and ss circles were made in
 vitro. Following HAP purification, this DNA was used as
 tracer in a subtractive hybridization reaction. The direct

FEATURES
 source

BASE COUNT 112 a 97 c 104 g 140 t
 ORIGIN
 Query Match 2.28; Score 34; EB 9; Length 4
 Best Local Similarity 100.0%; Pred. No. E=0.06;
 Matches 34; Conservative 0; Mismatches 0; Indels 0;

OY 1479 TTGCTCTTCTCAGTGTGACCTGATTAATAA 1512
 DB 289 TTGCTCTTCTCAGTGTGACCTGATTAATAA 22

RESULT 30
AA669250/1
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
SOURCE

AA669250 472 bp mRNA linear EST 20-NOV-1997
ac17c10.s1 Stralagene ovary (#937217) Homo sapiens cDNA clone
IMAGE:856746.3 similar to cclatlas Alu repetitive element, mRNA
sequence.

AA669250
EST
EST
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 472)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S.,
Kilman, D., Kudaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin,
J., Moore, B., Schellenger, K., Shapiro, M., Tan, F., Theisling, R.,
White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST project
Unpublished (1997)
Contact: Wilson, R.
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estw@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reversion of the 100% perfect match found
Seq primer: -40mt fwd, 5' from Amersham.
Location/Qualifiers
1. 472
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:856746"
/sex="female"
/dev_stage="49 year old"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ ovary, Vector: Plasmid SK, Site: 1. EcoRI
Site: 2. XbaI, cloned unidirectionally. Primer: oligo di-
total ovary tissue, normal, Caucasian. Average insert
size: 0.8 kb; Uni ZAP XR Vector; 5' adaptor sequence: 5'
CAATCGGCGGAG 3' -37 adaptor sequence: 5'
CTCGAGTTTCTTTTCTTTT 37"

BASE COUNT 124 a 110 c 93 g 145 t
ORIGIN

Query Match 1.7% Score 26; DB 9; Length 472;
Best local similarity 100.0%; Pred No. 0.092;
Matches 26; Dissimilative 0; Mismatches 0; Indels 0; Gaps 0

QY 1474 TTCATTGGTCTTCTGCTGTA 1499
|||||
Db 219 TTCATTGGCTGTTCTGAGTGA 194

RESULT 31
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
SOURCE

H00237 478 bp mRNA linear EST 19 JUN 1995
y322b03.r1 Soares placenta NB2HP Homo sapiens cDNA clone
IMAGE:149525.5, mRNA sequence.

H00237
EST
EST
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 478)
Hillier, L., Clark, N., Dubuque, T., Ellison, K., Haskins, M., Homan

AA669250/1
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
SOURCE

AA669250 472 bp mRNA linear EST 20-NOV-1997
ac17c10.s1 Stralagene ovary (#937217) Homo sapiens cDNA clone
IMAGE:856746.3 similar to cclatlas Alu repetitive element, mRNA
sequence.

AA669250
EST
EST
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 472)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S.,
Kilman, D., Kudaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin,
J., Moore, B., Schellenger, K., Shapiro, M., Tan, F., Theisling, R.,
White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST project
Unpublished (1997)
Contact: Wilson, R.
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estw@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reversion of the 100% perfect match found
Seq primer: -40mt fwd, 5' from Amersham.
Location/Qualifiers
1. 472
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:856746"
/sex="female"
/dev_stage="49 year old"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ ovary, Vector: Plasmid SK, Site: 1. EcoRI
Site: 2. XbaI, cloned unidirectionally. Primer: oligo di-
total ovary tissue, normal, Caucasian. Average insert
size: 0.8 kb; Uni ZAP XR Vector; 5' adaptor sequence: 5'
CAATCGGCGGAG 3' -37 adaptor sequence: 5'
CTCGAGTTTCTTTTCTTTT 37"

BASE COUNT 124 a 110 c 93 g 145 t
ORIGIN

Query Match 1.7% Score 26; DB 9; Length 472;
Best local similarity 100.0%; Pred No. 0.092;
Matches 26; Dissimilative 0; Mismatches 0; Indels 0; Gaps 0

QY 1474 TTCATTGGTCTTCTGCTGTA 1499
|||||
Db 219 TTCATTGGCTGTTCTGAGTGA 194

RESULT 32
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
SOURCE

AA669250 472 bp mRNA linear EST 20-NOV-1997
ac17c10.s1 Stralagene ovary (#937217) Homo sapiens cDNA clone
IMAGE:856746.3 similar to cclatlas Alu repetitive element, mRNA
sequence.

AA669250
EST
EST
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 478)
Hillier, L., Clark, N., Dubuque, T., Ellison, K., Haskins, M., Homan

Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multiplexedly sequenced cDNAs
10 (11), 1757-1771 (2000)
Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carrington, P., Sugahara,
Y., and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia. Real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamazawa, I., Aizawa,
F., Fukuoka, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., and
Hayashizaki, Y.
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome-gsc.riken.go.jp/>) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

FEATURES

Source

Location/Qualifiers
1..655
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="q1000418"
/clone_lib="RIKEN full length-enriched adult male germ"
/sex="male"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Site 1: SalI, site 2: BamHI. cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer (5'
GAGAGACGACGATTCGACGATCTTTTCTTTTCTTTCVN 3'). cDNA was
prepared by using trehalose thermo-stabilized reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rho - 10.0 and subtraction to Rho = 185.0. Second
strand cDNA was prepared with the primer adapter of
sequence (5' GAGAGACGATTCGACGATTAATAATACCCGCCCCCCC
3'). cDNA was cloned into the XhoI and BamHI sites.
Vector: a modified pRiscrypt KS(+) after bulk excision
from Lambda E1c I. Cloning sites, 5' end: SalI; 3' end:
BamHI"

BASE COUNT 176 a 160 c 165 g 153 t 1 others

Query Match 1.5%, Score 23, DB 1c, Length 655,
Best Local Similarity 100.0%, Pred. No. 2.9;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1476 GCTTCACCTCTCTGACCGACTT 1498
|||||
Db 613 GGTTCACCTCTCTGACCGACTT 635

RESULT 35
A1208287 417 bp mRNA linear EST 30-NOV-1998
Locus 995911.x1 Scares_testis_NHT Homo sapiens cDNA clone IMAGE:1839524
DEFINITION 3', mRNA sequence.
ACCESSION A1208287
VERSION A1208287.1 GI:3773229
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
ADDITIONS
TITLE
JOURNAL
COMMENT
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 417)
S1. <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: strauss@nci.nih.gov
cDNA library preparation: M. Bento Soares, Ph.D., M. Fatima Ronaldo
Ph.D.

cDNA library Arrayed by: Greg Lennon, Ph.D.
cDNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the 1 M A G E Consortium/1MAG at:
www.hioi.gov/hioi/mag/mag.html
Insert length: 523 Std Error: 0.00
Seq primer: 400P from Gibco.

FEATURES

Source

Location/Qualifiers
1..417
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1839524"
/clone_lib="Scares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pTZ19-pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI. 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I oligo(dT) primer (5'
TCTTACCAATTCGACGATGAGGACGACGACGACGATTCCTTTCTTTTCTTTCVN 3').
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pTZ19 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Ronaldo.

BASE COUNT 113 a 77 c 104 g 123 t

Query Match 1.4%, Score 22, DB 9; Length 417;
Best Local Similarity 100.0%, Pred. No. 9.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1477 GATTTGCTTTCTGAGTTGT 1498
|||||
Db 394 GATTTGCTTTCTGAGTTGT 373

RESULT 36
A0433394/c 555 bp DNA linear GSS 31-MAY-1999
Locus HS_5050_H1_F08_77A_R0C1-11 Human Male HAC Library Homo sapiens
ACCESSION A0433394
VERSION A0433394.1 GI:4543729
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 555)
AUTHORS Mahairas, G.G., Wallace, J.C., Smith, K., Swartz, L.L., Holzman, T.,
Keller, A., Shaker, K., Furlong, J., Young, J., Zhao, S., Adams, M.D., and
Hood, L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome.
Proc Natl Acad Sci U S A 96 (17), 9739-9744 (1999)
99380589
CONTACT: Mahairas G.G., Wallace J.C., Hood L.
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887

Email: jwallace@u.washington.edu
 Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Research Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu

Plate: 626 row: 1 column: 15
 Seq primer: 17
 Class: BAC ends
 High quality sequence stop: 555.
 Location/Qualifiers

1..555
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="plate=626 Col=15 Row=1"
 /clone_lib="RPCI-11 Human Male BAC Library"
 /sex="male"

/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at EcoRI sites"
 BASE COUNT 155 a 136 c 118 g 139 t 7 others
 ORIGIN

Query Match 1.4% Score 22; DB 17; Length 555.
 Best Local Similarity 100.0%; Pred. No. 9.4;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1478 ATTGGCTGTTCTGACTTGA 1499
 |||||||
 DB 494 ATTGGCTGTTCTGACTTGA 473

RESULT 37 629 bp DNA linear GSS 21-MAY-1999
 LOCUS RPC11-167M9 TV RPC1-11 Homo sapiens genomic clone RPC-11-167M9,
 DEFINITION DNA sequence.
 ACCESSION AC081994
 VERSION AC081994.1 GI:4354017
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 629)
 AUTHORS Zhao, S., Adams, M.D., Niernan, W., Malek, J., de Jong, P. and Venter, J.C.
 TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building
 JOURNAL Unpublished (1997)
 COMMENT Other GSS: RPC11-167M9.TJ
 Contact: Shaying Zhao, William Niernan, Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: hbe@ligr.org

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tigr/humgen/bac_end_search/bac_end_search.html
 Seq primer: 17
 Class: BAC ends.
 Location/Qualifiers

1..629
 /organism="Homo sapiens"
 /db_xref="GDB:7564040"

/db_xref="taxon:9606"

/clone="RPCI-11-167M9"

/clone_lib="RPCI-11"

/sex="Male"

/cell_type="lymphocytes"

/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; Human Male BAC Library"

BASE COUNT 177 a 123 c 146 g 182 t 1 others
 ORIGIN

Query Match 1.4% Score 22; DB 17; Length 555.
 Best Local Similarity 100.0%; Pred. No. 9.4;
 Matches 22; Conservative 0; Mismatches 0; Indels 0;

OY 1474 TTCATTTGGCTGTTCTGACT 1495
 |||||||
 DB 382 TTCATTTGGCTGTTCTGACT 403

RESULT 38 659 bp MPNA linear GSS 20-MAY-1999
 LOCUS BG390958
 DEFINITION 602417635F1 NIH_MGC_92 Homo sapiens cDNA clone H1
 mRNA sequence.
 ACCESSION BG390958
 VERSION BG390958.1 GI:13284406
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae;
 REFERENCE 1 (bases 1 to 659)
 AUTHORS NIH-MGC http://mgc.nhl.nih.gov/
 JOURNAL National Institutes of Health, Mammalian Gene
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-research.nhl.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consorti
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution info
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM10461 row: 1 column: 10
 High quality sequence stop: 659.
 Location/Qualifiers

1..659
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4536669"
 /clone_lib="NIH_MGC_92"
 /tissue_type="embryonal carcinoma, cell
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: testis; Vector: pCMV-Sport
 /site_2: SalI; Cloned unidirectionally;
 Average insert size 2.5 kb; library entry
 full length clones and constructed by 17
 Note: This is a NIH_MGC Library."

BASE COUNT 100 a 182 c 200 g 111 t
 ORIGIN

Query Match 1.4% Score 22; DB 12; Length 555.
 Best Local Similarity 100.0%; Pred. No. 9.4;
 Matches 22; Conservative 0; Mismatches 0; Indels 0;

OY 234 TCCAGCAATGAGGCTGCTGAC 315
 |||||||
 DB 538 TCCAGCAATGAGGCTGCTGAC 559

RESULT 39
 AL542749

```

LOCUS       AL542749                685 bp    mRNA    linear   EST_21 FEB 2001
DEFINITION  AL542749 LIT FL002_Pt1 Homo sapiens cDNA clone CS0E011YD05 5 prime
            A542749.1 G1:12875099
VERSION     AL542749.1
XREF        GenBank:AL542749
KEYWORDS    mRNA sequence.
ORGANISM    human.
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Cinclada; Vertebrata; Platyhelminthes;
            Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE   1 (bases 1 to 685)
AUTHORS     Li W.B., Gruber,C., Jessee,T and Polayes,D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 Evry cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
FEATURES             Source
                     1..685
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="CS0E011YD05"
                        /cologne_lib="LIT.FL002_Pt1"
                        /lab_host="DH10B"
                        /note="Organ: placenta, Vector: pCWSpcori 6' 1st strand
                        cDNA was primed with a NotI-oligo(dt) primer. Five prime
                        end enriched, double-stranded cDNA was digested with Not I
                        and cloned into the Not I and Eco RV sites of the
                        pCMVSPORT 6 vector. Library was constructed by life
                        technologies. Contact : Peng Liang life technologies, a
                        division of invitrogen 9800 Medical Center Drive Rockville
                        , Maryland 20850, USA Fax : (1) 301 610 8711 Email :
                        tliang@lifetechn.com npl :
                        http://fulllength.invitrogen.com"
BASE COUNT      209 a      154 c      160 g      158 t      4 others
ORIGIN
Query Match          1.4%; Score 22; DB 9; Length 685;
Best Local Similarity 100.0%; Pred. No. 9.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1477 CATTGAGCTTTCGATGTG 1498
DB 67 CATTGGCTGTTCTCATGTTG 98
|||||
RESULT 40
BF675057/c 822 bf MRNA linear EST_21 DEC 2000
ACCESSION BF675057
VERSION BF675057
KEYWORDS  NIH-MGC
SOURCE   EST.
ORGANISM human.
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Cinclada; Vertebrata; Platyhelminthes;
            Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
            1 (bases 1 to 822)
            NIH-MGC http://mgc.nhl.nih.gov/
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgabbs@mail.nih.gov
            Tissue Procurement: CHORI/IBCH Laboratories, Inc.
            cDNA Library Preparation: CHORI/IBCH Laboratories, Inc.
            DNA Library Arrayed by: The J.M.A.G.E. Consortium (JMNE)
            DNA sequencing by: Incyte Genomics, Inc.
            Clone distribution: Map clone distribution information can be
            found through the J.M.A.G.E. Consortium/JMNE at:
            http://image.jhu.edu
            Plate: LLCMI088 row: e column: 14

```

[illegible]

BASE COUNT 308 a 226 c 247 g 287 t 3 others
ORIGIN

Query Match 1.4% Score 22; DB 13; Length 1071;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 22; Conservative 0; Mismatches 0; Gaps 0;

UY 1477 CATTGGCTGTTCTGACTGCT 1498
|||||
DB 116 CATTGGCTGTTCTGACTGCT 137

RESULT 42

A1602007/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

210 bp mRNA linear EST 21-APR-1999
01-R-Y0-uv-q-02-0-01 s2 01-R-Y0 Rattus norvegicus cDNA clone
A1602007
A1602007 1 GI:4611168
EST.
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 210)
Honado, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.wppc.uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:
clones will be available through Research Genetics (www.resgen.com)
Seq primer: M13 Forward.
Location/Qualifiers
1..210
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="01-R-Y0-uv-q-02-0-01"
/clone_lib="01-R-Y0"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73B-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The 01-R-Y0
library is a subtracted library derived from at
individually-tagged normalized whole-eye (minus the lens)
library. The driver for the subtraction consisted of a
pool of all previous libraries (01-R-R0, 01-R-P1, 01-R-E0,
01-R-E1, 01-R-C0, and 01-R-C1). The tag is a string of
3-5 nucleotides present between the Not I site and the
oligo-dT track which allows identification of the library
of origin of a clone within the mixture. The subtracted
library (01-R-Y0) was constructed as follows: PCR
amplified cDNA inserts from previous library clones from
which 3' ESTs had been derived were used as a driver in a
hybridization with the normalized whole-eye library in
the form of single-stranded circles. The remaining
single-stranded circles (subtracted library) was purified
by hydroxyapatite column chromatography, converted to
double-stranded circles and electroporated into DH10B
bacteria (Life Technologies) to generate the 01-R-Y0

library. This procedure has been previously
(Bonafide, Lennon and Soares, Genome Res.
1996)"
BASE COUNT 55 a 41 c 44 g 70 t
ORIGIN

Query Match 1.4% Score 21; DB 9; Length 21;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 21; Conservative 0; Mismatches 0; Gaps 0;

CY 1522 TGTGTGACGCAAAAAAAA 1542
|||||
DB 29 TGTGTGACGCAAAAAAAA 9

RESULT 43

A0270139/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

277 bp DNA linear
HS-2060_A1_H12_T7 c11 Approved Human Genomic Spec
sapient genomic clone Plate-2060 Col-23 Row-0, INT
A0270139
A0270139.1 GI:3824204
GSS.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata;
Mammalia; Eutheria; Primates; Catarrhini; Homiida;
1 (bases 1 to 377)
Mahairas, G.G., Wallace, J.C., Smith, K., Spartz, J.,
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, J.,
Hood, L.
Sequence-tagged connectors: A sequence approach to
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744
99380589
Contact: Mahairas G.G. Wallace J.C., Hood L.
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, US
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: wallace@u.washington.edu
Sequence Tagged Connector
Plate: 2060 Row: 0 Column: 23
Class: BAC ends
High quality sequence stop: 477.
Location/Qualifiers
1..377
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate 2060 Col-23 Row-0"
/clone_lib="c11 Approved Human Genomic Sp
/sex="male"
/note="Organ: Sperm; Vector: pBluescript II
E-Coll DH10B"
BASE COUNT 99 a 102 c 91 t 3 others
ORIGIN

Query Match 1.4% Score 21; DB 17; Length 4;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 21; Conservative 0; Mismatches 0; Gaps 0;

UY 1406 CCACAGAGCGATGAGCGTC 1426
|||||
DB 236 CCACAGAGCGATGAGCGTC 216

RESULT 44

A1764709/c

LOCUS

DEFINITION

ACCESSION

404 bp mRNA linear
01-R-Y0-ace-b-07-0-01 s1 01-R-Y0 Rattus norvegicus
01-R-Y0-ace-b-07-0-01 s1, mRNA sequence.
A1764709

VERSION A1764704.1 G1:5210644
 KEYWORDS EST.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 404)
 AUTHORS Ronalds, M.F., Lennon, G., and Soares, M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 COMMENT 97044477
 Contact: Soares, MR
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@uiowa.edu
 The sequence contained an oligo-dT track that was present in the
 oligonucleotide that was used to prime the synthesis of first
 strand cDNA and therefore this may represent a bonafide poly A
 tail. The sequence tag present in the cDNA between the NotI site
 and the oligo-dT track served to verify it as a clone from the
 normalized eye library cDNA library preparation; M.B. Soares Lab
 Clone distribution: clones will be available through Research
 Genetics (www.resgen.com)
 Seq primer: M13 Forward
 POLYA-Yes.
 FEATURES
 SOURCE location/qualifiers
 1 404
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="U1-R-Y0-300-b-07-0-01"
 /clone_lib="U1-R-Y0"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pTZ19-Pac (Pharmacia) with a modified
 polylinker site. Not I site 2. Eco RI. The U1-R-Y0
 library is a subtracted library derived from an
 individually-tagged normalized whole-eye (minus the lens)
 library. The driver for the subtraction consisted of a
 pool of all previous libraries (U1-R-A0, U1-R-A1, U1-R-E0,
 U1-R-E1, U1-R-E0, and U1-R-G1). The tag is a string of
 3-5 nucleotides present between the Not I site and the
 oligo-dT track which allows identification of the library
 of origin of a clone within the mixture. The subtracted
 library (U1-R-Y0) was constructed as follows: pre-
 amplified cDNA inserts from previous library clones for
 which 3' ESTs had been derived were used as a driver in a
 hybridization with the normalized whole eye library in
 the form of single-stranded circles. The remaining
 single-stranded circles (subtracted library) was purified
 by hydroxyapatite column chromatography, converted to
 double-stranded circles and electroporated into DH10B
 bacteria (Life Technologies) to generate the U1-R-Y0
 library. This procedure has been previously described
 (Bonaldi, Lennon and Soares, Genome Research 6: 791-806,
 1996)
 TAG_LIB=U1-R-Y0
 TAG_TISSUE=eye
 TAG_SEQ=CATTG*
 BASE COUNT 93 a 105 c 95 g 111 t
 ORIGIN
 Query Match 1.4% Score 21; DB 9; Length 404;
 Best Local Similarity 100.0%; Prod. No. 31;
 Matches 21; Conservative 0; Mismatches 0; Gaps 0;
 OY 1522 TCTCTACGCGCAAAAAA 1542
 ||||||||||||||||||||

DB 29 TCTCTACGCGCAAAAAA 9
 RESULT 45
 AC663408 431 bp DNA linear GSS 23-JUN-1999
 LOCUS Hs_216042.109-270 c17 Approved Human Genomic Sperm Library D Homo
 DEFINITION sapiens genomic clone Plate 2160 Col-12 Row-H, DNA sequence.
 ACCESSION AC663408
 VERSION AC663408.1 G1:5171176
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 431)
 AUTHORS Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holman, T.,
 Kolter, A., Stokoe, R., Fortington, J., Young, T., Zhao, S., Adams, M.D., and
 Hood, L.
 TITLE Sequence-tagged cDNAs: A sequence approach to mapping and
 scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 COMMENT Contact: Mahairas GG, Wallace JG, Hood L.
 High Throughput Sequencing Center
 University of Washington
 451 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones may be purchased from Research Genetics (info@resgen.com),
 RGC and Web Server: <http://www.brsc.washington.edu>
 Plate: 2160 Row: H Column: 12
 Seq primer: T7
 Class: BAC ends
 High quality sequence stop: 431
 FEATURES
 SOURCE location/qualifiers
 1 431
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="H1300-2160 Col 12 Row-H"
 /clone_lib="C17 Approved Human Genomic Sperm Library D"
 /sex="male"
 /note="organ: sperm; vector: pTZ19AC11; BAC clones in
 E-Coli DH10B"
 BASE COUNT 124 a 101 c 116 g 1 others
 ORIGIN
 Query Match 1.4% Score 21; DB 17; Length 431;
 Best Local Similarity 100.0%; Prod. No. 31;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 GY 1477 CATTGCTGTCTGACTTC 1497
 ||||||||||||||||||||
 DB 227 CATTGCTGTCTGACTTC 247
 Search completed: November 7, 2002, 11:02:15
 Job time: 2098 secs

Q7 1472 TCTTCATTGGCTGTTCTGA 1493
|||||
Db 64 TCTTCATTGGCTGTTCTGA 42

RESULT 7
US-09-604-287A-29/C
Sequence 29, Application US/09604287A
Patent No. US20020064872A1
GENERAL INFORMATION:
APPLICANT: Jiang, Yugu
APPLICANT: Dillon, David C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jiangchun
APPLICANT: Harlocker, Susan L.
APPLICANT: Hepler, William T.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.47007
CURRENT FILING DATE: 2000-06-22
NUMBER OF SEQ ID NOS: 489
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 29
LENGTH: 301
TYPE: DNA
ORGANISM: Homo sapien
US-09-604-287A-29

Query Match 1.28, Score 19, DB 10, Length 301;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 19, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

Q7 1060 GTGCAATGGCTGACGGC 1078
|||||
Db 265 GTGCAATGGCTGACGGC 247

RESULT 8
US-09-339-338-29/C
Sequence 29, Application US/09339338A
Patent No. US20020102602A1
GENERAL INFORMATION:
APPLICANT: Yugu, Jiang
APPLICANT: Dillon, David C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jiangchun
TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
FILE REFERENCE: 210121.47007
CURRENT APPLICATION NUMBER: US/09/458A
CURRENT FILING DATE: 1999-05-24
NUMBER OF SEQ ID NOS: 315
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 29
LENGTH: 301
TYPE: DNA
ORGANISM: Homo sapien
US-09-339-338-29

Query Match 1.28, Score 19, DB 10, Length 301;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 19, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

Q7 1060 GTGCAATGGCTGACGGC 1078
|||||
Db 265 GTGCAATGGCTGACGGC 247

RESULT 9
US-10-007-805-29/C
Sequence 29, Application US/10007805

Patent No. US20020150581A1
GENERAL INFORMATION:
APPLICANT: Jiang, Yugu
APPLICANT: Dillon, David C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jiangchun
APPLICANT: Harlocker, Susan L.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Fung, Gary K.
APPLICANT: Vedvick, Thomas S.
APPLICANT: McNeill, Patricia D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.470010
CURRENT APPLICATION NUMBER: US/10/007,805
CURRENT FILING DATE: 2001-12-07
NUMBER OF SEQ ID NOS: 593
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 29
LENGTH: 301
TYPE: DNA
ORGANISM: Homo sapien
US-10-007-805-29

Query Match 1.28, Score 19, DB 12, Length 301;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 19, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

Q7 1060 GTGCAATGGCTGACGGC 1078
|||||
Db 245 GTGCAATGGCTGACGGC 247

RESULT 10
US 09 864 761 1342
Sequence 1342, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Kim, Sharon C.
APPLICANT: Hanzel, David R.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Neomiga-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 96/190,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 96/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 99/632,366
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/2246,459
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 621,0001/000666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 621,0001/000667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 621,0001/000664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 621,0001/000659
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 621,0001/000655
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 621,0001/000658
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 621,0001/000663
PRIOR FILING DATE: 2001-01-30

```

PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annonax Sequence Listing Engine v0rs. 1.1
SEQ ID NO: 13242
LENGTH: 520
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC017079.2
OTHER INFORMATION: EXPRESSED IN HPAIN, SIGNAL - 1
US-09-854-761-13242

```

```

Query Match          12%: Score 19, DB 10, Length 520:
Best Local Similarity 100.0%: Pred. No. 7.1,
Matches 19, Conservative 0, Mismatches 0, Indels 0, Gaps 0

```

```

QY 474 AACGTGAGTGGTGAGCT 492
    |||||
DB 169 AACGTGAGTGGTGAGCT 187

```

RESULT 11

```

US-09-905-291A-376
Sequence 376, Application US/09905291A
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Bolstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavitt, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William L.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/905,291A
PRIOR FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26

```

```

PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-14
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/20809
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/20214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28413
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28595
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30929
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO: 376
LENGTH: 997
TYPE: DNA
ORGANISM: Homo Sapien
US-09-905-291A-376

```

```

Query Match          12%: Score 19, DB 9, Length 997
Best Local Similarity 100.0%: Pred. No. 7.3,
Matches 19, Conservative 0, Mismatches 0, Indels 0, Gaps 0

```

```

QY 1060 GTGGCAATGGTGAGCGGC 1078
    |||||
DB 552 GTGGCAATGGTGAGCGGC 570

```

RESULT 12

```

US-09-909-320-376
Sequence 376, Application US/09909320
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Bolstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavitt, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey

```

```

APPLICANT: Wood, William, L.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/099 320
CURRENT FILING DATE: 2002-01-04
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21065
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28413
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/29555
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 376
LENGTH: 997
TYPE: DNA
ORGANISM: Homo Sapien
US-09-909-320-376

Query Match: 1 26, Score 13; DB 10; Length 997;
Best Local Similarity 100.0%; Pred No. 7.3;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1060 GTGGCAATGGGTGACGGC 1078
|||||
Db 552 GTGGCAATGGGTGACGGC 570

RESULT 13
US-09-909-088B-376
Sequence 376, Application US/09909/088B
Patent No. US20020146709A1
GENERAL INFORMATION:
APPLICANT: Genentech, Inc
APPLICANT: Ashkenazi, Avi
APPLICANT: Rotstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eyalon, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroli, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Garber, Hanspeter
APPLICANT: Gertlisen, Mary E.
APPLICANT: Goddard, A.

```

```

APPLICANT: Rodowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kistritz, Ivor J.
APPLICANT: Malher, Jennie E.
APPLICANT: Pao, James
APPLICANT: Pao, Nicholas E.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Thomas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, L.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/099 088B
CURRENT FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21593
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28314
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 376
LENGTH: 997
TYPE: DNA
ORGANISM: Homo Sapien
US-09-909-088B-376

Query Match: 1 26, Score 13; DB 10; Length 997;
Best Local Similarity 100.0%; Pred No. 7.3;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1060 GTGGCAATGGGTGACGGC 1078
|||||
Db 552 GTGGCAATGGGTGACGGC 570

RESULT 14
US-10-052-586-23
Sequence 23, Application US/10052586
Patent No. US20020127584A1
GENERAL INFORMATION:

```

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Han, James
APPLICANT: Smith, Victoria K.
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C1
CURRENT APPLICATION NUMBER: US/10/052,586
CURRENT FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062550
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063564
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063714
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063870
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066120
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066456
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066772
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069345
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069425
PRIOR FILING DATE: 1997-12-12
PRIOR APPLICATION NUMBER: 60/069870
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/068017
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/074510
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077642
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/078446
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078449
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080107

PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080104
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080343
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081848
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/083422
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084466
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084644
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084644
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085802
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085802
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/085809
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087779
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087708
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087779
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04


```

: TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
: TITLE OF INVENTION: Sels
: FILE REFERENCE: 689290-76
: CURRENT APPLICATION NUMBER: US/09/954,456
: CURRENT FILING DATE: 2001-09-18
: PRIOR APPLICATION NUMBER: US/60/233,617
: PRIOR FILING DATE: 2000-09-18
: PRIOR APPLICATION NUMBER: US/60/234,052
: PRIOR FILING DATE: 2000-09-20
: PRIOR APPLICATION NUMBER: US/60/234,923
: PRIOR FILING DATE: 2000-09-25
: PRIOR APPLICATION NUMBER: US/60/235,134
: PRIOR FILING DATE: 2000-09-25
: PRIOR APPLICATION NUMBER: US/60/235,637
: PRIOR FILING DATE: 2000-09-26
: PRIOR APPLICATION NUMBER: US/60/235,638
: PRIOR FILING DATE: 2000-09-26
: PRIOR APPLICATION NUMBER: US/60/235,711
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,720
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,840
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,863
: PRIOR FILING DATE: 2000-09-27
: NUMBER OF SEQ ID NOS: 2276
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO: 1685
: LENGTH: 456
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-954-456-1685

```

```

Query Match          1.28; Score 18; DB 10; Length 456.
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 503 TCCAGCTGCAGGAGATC 520
    |||||||
DB 81 TCCAGCTGCAGGAGATC 64

```

```

RESULT 17
US-10-044-090-690/c
: Sequence 690, Application US/10044090
: Patent No. US20020137081A1
: GENERAL INFORMATION:
: APPLICANT: Olga Bandman
: TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
: FILE REFERENCE: PA-002A US
: CURRENT APPLICATION NUMBER: US/10/044,090
: CURRENT FILING DATE: 2002-01-09
: NUMBER OF SEQ ID NOS: 850
: SOFTWARE: PERL Program
: SEQ ID NO: 690
: LENGTH: 1389
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incycle ID No. US20020137081A1 480287.8
US-10-044-090-690

```

```

Query Match          1.28; Score 18; DB 12; Length 1389;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 158 TTCTGTTGTTGTTGCTG 175
    |||||||
DB 964 TTCTGTTGTTGTTGCTG 947

```

RESULT 18

```

US-09-729-674-19
: Sequence 19, Application US/09729674
: Patent No. US20010039335A1
: GENERAL INFORMATION:
: APPLICANT: Jacobs, Kenneth
: APPLICANT: McCoy, John M.
: APPLICANT: Lavallee, Edward E.
: APPLICANT: Collins-Racie, Lisa A.
: APPLICANT: Evans, Cheryl
: APPLICANT: Merberg, David
: APPLICANT: Treacy, Maurice
: APPLICANT: Apostoluo, Michael J.
: APPLICANT: Steinhilber II, Robert J.
: APPLICANT: Spaulding, Vikki
: APPLICANT: Wong, Gordon G.
: APPLICANT: Clark, Hilary
: APPLICANT: Fechtel, Kim
: APPLICANT: Genetics Institute, Inc.
: TITLE OF INVENTION: SECRETED PROTEINS AND POLYPEPTIDES
: FILE REFERENCE: 6055-64X
: CURRENT APPLICATION NUMBER: US/09/729,674
: CURRENT FILING DATE: 2000-12-04
: PRIOR APPLICATION NUMBER: US/75/9,109
: PRIOR FILING DATE: 2000-03-30
: NUMBER OF SEQ ID NOS: 283
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO: 19
: LENGTH: 1656
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-729-674-19

```

```

Query Match          1.28; Score 18; DB 10; Length 1656
Best Local Similarity 100.0%; Pred. No. 24;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1525 TGTACTGCAAAAAAAA 1542
    |||||||
DB 1596 TGTACTGCAAAAAAAA 1613

```

```

RESULT 19
US-09-764-869-2142/c
: Sequence 2142, Application US/09764869
: Patent No. US20020061521A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodi
: FILE REFERENCE: PC007
: CURRENT APPLICATION NUMBER: US/09/764,869
: PRIOR FILING DATE: 2001-01-17
: PRIOR APPLICATION data removed - refer to PA1M or file with
: NUMBER OF SEQ ID NOS: 2442
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO: 2142
: LENGTH: 6800
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-764-869-2142

```

```

Query Match          1.28; Score 18; DB 10; Length 6800
Best Local Similarity 100.0%; Pred. No. 24;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 991 ACAGCAGCAGACTTCAG 1008
    |||||||
DB 1443 ACAGCAGCAGACTTCAG 1426

```

```

RESULT 20
US-09-917-800A-502
: Sequence 502, Application US/09917800A
: Patent No. US20020119462A1

```

```

GENERAL INFORMATION:
APPLICANT: Mendick, Donna
APPLICANT: Porter, Mark
APPLICANT: Johnson, Kory
APPLICANT: Castle, Arthur
APPLICANT: Elashoff, Michael
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Molecular Toxicology Modeling
FILE REFERENCE: 44921-5038-US
CURRENT APPLICATION NUMBER: US/09/917,800A
CURRENT FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: US 60/222,040
PRIOR FILING DATE: 2000-07-31
PRIOR APPLICATION NUMBER: US 60/222,880
PRIOR FILING DATE: 2000-11-02
PRIOR APPLICATION NUMBER: US 60/240,029
PRIOR FILING DATE: 2001-05-11
PRIOR APPLICATION NUMBER: US 60/250,645
PRIOR FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: US 60/292,336
PRIOR FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US 60/295,798
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US 60/297,457
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,984
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/303,459
PRIOR FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 1740
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 502
LENGTH: 7420
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURES:
OTHER INFORMATION: Genbank Accession No. US20020114462A1 Archived
US-09-917-800A-502

Query Match
Best Local Similarity 100.0%; Score 18; DB 10; Length 7420
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 450 TCTACGAGACCTGACCC 467
|||||
DB 7286 TCTACGAGACCTGACCC 7303

RESULT 21
US-09-954-456-2006
Sequence 2006, Application US/09/954,456
Patent No. US20020115057A1
GENERAL INFORMATION:
APPLICANT: Young, Paul
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
FILE REFERENCE: 689290-76
CURRENT APPLICATION NUMBER US/09/954,456
CURRENT FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/60/233,617
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,052
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,923
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,134
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,637
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,648
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,711
PRIOR FILING DATE: 2000-09-27

```

```

PRIOR APPLICATION NUMBER: US/60/235,720
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,840
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,863
PRIOR FILING DATE: 2000-09-27
NUMBER OF SEQ ID NOS: 2276
SOFTWARE: Patent In version 4.0
SEQ ID NO 2006
LENGTH: 7787
TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-456-2006

Query Match
Best Local Similarity 100.0%; Score 18; DB 10; Length 7787;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 450 TCTACGAGACCTGACCC 467
|||||
DB 7388 TCTACGAGACCTGACCC 7405

RESULT 22
US-09-962-832-154
Sequence 154, Application US/09/962,832
Patent No. US2002010821A1
GENERAL INFORMATION:
APPLICANT: Ebner, Reinhard
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
FILE REFERENCE: 689290-74
CURRENT APPLICATION NUMBER: US/09/962,832
CURRENT FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US/60/235,077
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,290
PRIOR FILING DATE: 2000-09-25
NUMBER OF SEQ ID NOS: 259
SOFTWARE: Patent In version 4.0
SEQ ID NO 154
LENGTH: 302250
TYPE: DNA
ORGANISM: Homo sapiens
US-09-962-832-154

Query Match
Best Local Similarity 100.0%; Score 18; DB 10; Length 302250;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160218 CCCTTCGACGCGGCGACT 160235
|||||
DB 160218 CCCTTCGACGCGGCGACT 160235

RESULT 23
US-09-560-863-384
Sequence 384, Application US/09/560,863
Patent No. US20020110809A1
GENERAL INFORMATION:
APPLICANT: Neffs, Michael C.
APPLICANT: Gashrowicz, Brian
TITLE OF INVENTION: No. US20020110809A1 Human Polynucleotides and the
FILE REFERENCE: LEX-0018-USA
CURRENT APPLICATION NUMBER: US/09/560,863
CURRENT FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/132,404
PRIOR FILING DATE: 1999-04-10
NUMBER OF SEQ ID NOS: 1008
SOFTWARE: Fastdb for Windows Version 4.0
SEQ ID NO 384

```

```

: LENGTH: 353
: TYPE: DNA
: ORGANISM: homo sapiens
: US-09-560-863-384

Query Match
Best Local Similarity 100.0%; Score 17; DB 10; Length 353;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1526 GTAACTGCAAAAAAAA 1542
Db 337 GTAACTGCAAAAAAAA 353

RESULT 24
US-09-864-761-2789/c
: Sequence 2789, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: FILE REFERENCE: Aeonica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263,6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 2789
: LENGTH: 381
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC009229.1
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.3
```

```

: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.5
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
: OTHER INFORMATION: EXPRESSED IN PL474, SIGNAL = 2.9
: OTHER INFORMATION: EXPRESSED IN HONE MARKER, SIGNAL = 4.0
: OTHER INFORMATION: EXPRESSED IN H8100, SIGNAL = 2.8
: OTHER INFORMATION: EXPRESSED IN H1AL, SIGNAL = 4.4
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.4
: OTHER INFORMATION: EXPRESSED IN H1LA, SIGNAL = 5.5
: OTHER INFORMATION: EXPRESSED IN H1AK, SIGNAL = 3.6
: US-09-864-761-2789

Query Match
Best Local Similarity 100.0%; Score 17; DB 10; Length
Matches 17; Conservative 0; Mismatches 0; Indels 0;

QY 382 GGGCATCTTGTTCAC 394
Db 348 GGGCATCTTGTTCAC 342

RESULT 25
US-09-783-590-7957
: Sequence 7957, Application US/09783590
: Patent No. US20020110850A1
: GENERAL INFORMATION:
: APPLICANT: Dillon, Patrick J.
: APPLICANT: Haseltine, William A.
: APPLICANT: Li, Haodong
: APPLICANT: Rosen, Craig A.
: APPLICANT: Ruben, Steven M.
: TITLE OF INVENTION: Human Genes, Sequences, and Expression
: FILE REFERENCE: PO-16,201
: CURRENT APPLICATION NUMBER: US/09/783,590
: CURRENT FILING DATE: 2000-02-15
: PRIOR APPLICATION NUMBER: 09/420,856
: PRIOR FILING DATE: 1995-04-12
: PRIOR APPLICATION NUMBER: 09/446,741
: PRIOR FILING DATE: 1994-11-21
: NUMBER OF SEQ ID NOS: 12485
: SOFTWARE: Patent In Ver. 2.0
: SEQ ID NO 7957
: LENGTH: 442
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: (10)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (26)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (189)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (203)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (286)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (313)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (316)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (327)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (338)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
```

```

LOCATION: (394)
OTHER INFORMATION: n equals a,t,q, or c
NAME/KEY: misc feature
LOCATION: (395)
OTHER INFORMATION: n equals a,t,q, or c
NAME/KEY: misc feature
LOCATION: (398)
OTHER INFORMATION: n equals a,t,q, or c
NAME/KEY: misc feature
LOCATION: (404)
OTHER INFORMATION: n equals a,t,q, or c
NAME/KEY: misc feature
LOCATION: (427)
OTHER INFORMATION: n equals a,t,q, or c
NAME/KEY: misc feature
LOCATION: (435)
OTHER INFORMATION: n equals a,t,q, or c
US-09-783-590-7957

```

```

Query Match 1.1% Score 17, DB 10, Length 412,
Best Local Similarity 100.0%, Pred. No. 74,
Matches 17, Conservative 0, Mismatches 0, Indels 0, Gaps 0

```

```

QY 866 TCTTCACGACGACG 882
DB 291 TCTTCACGACGACG 307

```

```

RESULT 26
US-09-864-761-1416/c
Sequence 1416, Application US/09864761
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David P.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DEPENDED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/432,366
PRIOR FILING DATE: 2000-04-03
PRIOR APPLICATION NUMBER: GR 24233.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,259
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00656
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00657
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00654
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00659
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687

```

```

PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 1416
LENGTH: 452
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO A002962.1
OTHER INFORMATION: EXPRESSED IN HELLO, SIGNAL = 2.6
OTHER INFORMATION: EXPRESSED IN HEAL, SIGNAL = 3.3
OTHER INFORMATION: EXPRESSED IN HEAL, SIGNAL = 3.9
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.4
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.6
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.3
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.2
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4
US-09-864-761-1416

```

```

Query Match 1.1% Score 17, DB 10, Length 452,
Best Local Similarity 100.0%, Pred. No. 74,
Matches 17, Conservative 0, Mismatches 0, Indels 0, Gaps 0

```

```

QY 1079 TCACCTTCGTCACGACG 1095
DB 242 TCACCTTCGTCACGACG 226

```

```

RESULT 27
US-09-974-300-2991
Sequence 2991, Application US/09974300
Patent No. US20020146721A1
GENERAL INFORMATION:
APPLICANT: Herko, Randy M.
APPLICANT: Clausen, ID Groth
TITLE OF INVENTION: Methods For Monitoring Multiple Gene
FILE REFERENCE: 10085,500-US
CURRENT APPLICATION NUMBER: US/09/974,300
CURRENT FILING DATE: 2001-10-05
PRIOR APPLICATION NUMBER: 09/440,538
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/275,526
PRIOR FILING DATE: 2001-03-27
NUMBER OF SEQ ID NOS: 8481
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2991
LENGTH: 573
TYPE: DNA
ORGANISM: Bacillus licheniformis
US-09-974-300-2991

```

```

Query Match 1.1% Score 17, DB 10, Length 573,
Best Local Similarity 100.0%, Pred. No. 75,
Matches 17, Conservative 0, Mismatches 0, Indels 0, Gaps 0

```

```

QY 591 ACACGACGACGACGACG 607
DB 245 ACACGACGACGACGACG 252

```

```

RESULT 28
US-09-864-761-17795/c
Sequence 17795, Application US/09864761
Patent No. US20020048753A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.

```

```

APPLICANT: Hank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GR 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 17795
LENGTH: 577
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC009229.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.3
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.5
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.9
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.4
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.8
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.4
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.5
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.6
OTHER INFORMATION: SWISSPROT HIT: Q16674, EVALUATE 1.00e-111
OTHER INFORMATION: NT HIT: Q111430100, EVALUATE 0.00e+00
OTHER INFORMATION: EST_HUMAN HIT: A0120297.1, EVALUATE 0.00e+00
US-09-864-761-17795

```

```

Query Match 1.18, Score 17, DH 10, Length 577;
Best Local Similarity 100.08; Pred. No. 75;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 382 GATGATCTTGTCCAG 398
|||||

```

```

Db 308 GGCATCTTGTCCAG 292
RESULT 29
US-09-864-761-16207
Sequence 16207, Application 05/09844761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Hank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GR 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 16207
LENGTH: 592
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO A116965.2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2
US-09-864-761-16207

```

```

Query Match 1.18, Score 17, DH 10, Length 592;
Best Local Similarity 100.08; Pred. No. 75;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 358 GATGCTTGTGCTGGG 374
|||||
Db 269 CATGCTTGTGCTGGG 285

```

```

Query      293 GTCAGATAGTGGGCT 309
Db         742 GTGATAGTGGGCT 726

RESULT 32
US-09-916-790-3
Sequence 3, Application US/09916790
Patent No. US20020061573A1
GENERAL INFORMATION:
APPLICANT: Meyers, Rachael
APPLICANT: Kapeller-Libermann, Rosana
TITLE OF INVENTION: 18431 AND 12374, NOVEL HUMAN PROTEIN
FILE REFERENCE: 381552002700
CURRENT APPLICATION NUMBER: US/09/916,790
PRIOR FILING DATE: 2001-03-27
PRIOR APPLICATION NUMBER: 60/221,543
NUMBER OF SEQ ID NOS: 37
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 1041
TYPE: DNA
ORGANISM: Homo sapiens
US-09-916-790-3

Query Match      1.1%, Score 17, DB 10, Length 1041;
Best Local Similarity 109.0%; Pred. No. 77;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      293 GTCAGATAGTGGGCT 309
DB      742 GTGATAGTGGGCT 726

RESULT 33
US-09-938-842A-134
Sequence 134, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harpoy, Jeff
APPLICANT: Kropf, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
TITLE OF INVENTION: SAME, AND METHODS OF USE
FILE REFERENCE: SCRIPT-300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
PRIOR FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 09/227,896
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 09/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 134
LENGTH: 1059
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-134

Query Match      1.1%, Score 17, DB 9, Length 1059;
Best Local Similarity 109.3%; Pred. No. 78;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1396 CCAATGCAATCCAGACA 1412
DB      742 GTGATAGTGGGCT 726

```

DB 828 CTAAGCATCTGAGAGCA 844

RESULT 34

US-09-833-381-71

Sequence 71, Application US/09833381

Patent No. US20020132090A1

GENERAL INFORMATION:

APPLICANT: Robison, Keith E.

TITLE OF INVENTION: NO. US20020132090A1: Nucleic Acid and Protein Homologs

FILE REFERENCE: 5800-119

CURRENT APPLICATION NUMBER: US/09/833,381

PRIOR FILING DATE: 2001-04-11

PRIOR APPLICATION NUMBER: 09/516,448

NUMBER OF SEQ ID NOS: 2050

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 71

LENGTH: 1154

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1) (1154)

OTHER INFORMATION: n - A,T,C or G

US-09-833-381-71

Query Match 1.18; Score 17; DB 10; Length 1154,
Best Local Similarity 100.0%; Pred. No. 78;

Matches 17; Conservative 0; Mismatches 0; Indels 0;

QY 243 TGAACATCTGAGAGCA 259

|||||

DB 269 TGAACATCTGAGAGCA 285

RESULT 35

US-09-917-800A-1377

Sequence 1377, Application US/09917800A

Patent No. US2002019462A1

GENERAL INFORMATION:

APPLICANT: Mendrick, Donna

APPLICANT: Porter, Mark

APPLICANT: Johnson, Kory

APPLICANT: Castle, Arthur

APPLICANT: Elashoff, Michael

TITLE OF INVENTION: Molecular Toxicology Modeling

FILE REFERENCE: 44921-5038-05

CURRENT APPLICATION NUMBER: US/09/917,800A

PRIOR FILING DATE: 2001-07-31

PRIOR APPLICATION NUMBER: US 60/222,040

PRIOR FILING DATE: 2000-07-31

PRIOR APPLICATION NUMBER: US 60/222,880

PRIOR FILING DATE: 2000-11-02

PRIOR APPLICATION NUMBER: US 60/290,029

PRIOR FILING DATE: 2001-05-11

PRIOR APPLICATION NUMBER: US 60/290,645

PRIOR FILING DATE: 2001-05-15

PRIOR APPLICATION NUMBER: US 60/292,336

PRIOR FILING DATE: 2001-05-22

PRIOR APPLICATION NUMBER: US 60/295,798

PRIOR FILING DATE: 2001-06-06

PRIOR APPLICATION NUMBER: US 60/297,457

PRIOR FILING DATE: 2001-06-13

PRIOR APPLICATION NUMBER: US 60/298,884

PRIOR FILING DATE: 2001-06-19

PRIOR APPLICATION NUMBER: US 60/303,459

PRIOR FILING DATE: 2001-07-04

NUMBER OF SEQ ID NOS: 1740

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 1377

LENGTH: 1164

TYPE: DNA

ORGANISM: Rattus norvegicus

FEATURE:

OTHER INFORMATION: Genbank Accession No. US20020119442A1

US-09-917-800A-1377

Query Match 1.18; Score 17; DB 10; Length
Best Local Similarity 100.0%; Pred. No. 78;

Matches 17; Conservative 0; Mismatches 0; Indels 0;

QY 1354 CTTCGAGCTGAGAGCA 1370

|||||

DB 992 CTTCGAGCTGAGAGCA 1008

RESULT 36

US-10-100-745-4

Sequence 3, Application US/10100745

Patent No. US20020115169A1

GENERAL INFORMATION:

APPLICANT: Olsson, Mark

APPLICANT: Szebenl, Atilla

TITLE OF INVENTION: METHODS FOR STABILIZATION AND

FILE REFERENCE: 13121.000203

CURRENT APPLICATION NUMBER: US/10/100,745

PRIOR FILING DATE: 2002-04-18

PRIOR APPLICATION NUMBER: 60/548,871

PRIOR FILING DATE: 2000-03-30

PRIOR APPLICATION NUMBER: 60/7126,910

PRIOR FILING DATE: 1999-03-30

NUMBER OF SEQ ID NOS: 4

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 3

LENGTH: 1164

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of artificial sequence

US-10-100-745-4

Query Match 1.18; Score 17; DB 12; Length 1
Best Local Similarity 100.0%; Pred. No. 78;

Matches 17; Conservative 0; Mismatches 0; Indels 0;

QY 1354 CTTCGAGCTGAGAGCA 1370

|||||

DB 992 CTTCGAGCTGAGAGCA 1008

RESULT 37

US-09-734-032-1

Sequence 1, Application US/09734032

Patent No. US20020103116A1

GENERAL INFORMATION:

APPLICANT: Wei et al

TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEI

FILE REFERENCE: C000536

CURRENT APPLICATION NUMBER: 60/207,442

PRIOR FILING DATE: 2001-08-16

PRIOR APPLICATION NUMBER: 60/207,228

PRIOR FILING DATE: 2000-05-17

NUMBER OF SEQ ID NOS: 4

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1

LENGTH: 1275

TYPE: DNA

ORGANISM: human

US-09-734-032-1

Query Match 1.1%; Score 17; DB 10; Length 1275;
 Best Local Similarity 100.0%; Pred. No. 78;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GTGCAGCAGCTGGGCGCT 309
 |||||
 Db 463 GTGCAGCAGCTGGGCGCT 479

RESULT 38
 US-10-016-985-1
 : Sequence 1, Application US/10016985
 : Patent No. US20020123621A1
 : GENERAL INFORMATION:
 : APPLICANT: Walke, D. Wade
 : APPLICANT: Maricar, Miranda
 : APPLICANT: Yu, Xuanhuan (Sean)
 : APPLICANT: Fiddie, Carl Johan
 : TITLE OF INVENTION: No. US-985-123621A1 Human Kinase and Polynucleotides
 : TITLE OF INVENTION: Encoding the Same
 : FILE REFERENCE: LEX-0273-USA
 : CURRENT APPLICATION NUMBER: US/10/016,985
 : PRIOR FILING DATE: 2001-12-07
 : PRIOR APPLICATION NUMBER: US 60/251,941
 : NUMBER OF SEQ ID NOS: 3
 : SOFTWARE: FASTSEQ for Windows Version 4.0
 : SEQ ID NO 1
 : LENGTH: 1275
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 US-10-016-985-1

Query Match 1.1%; Score 17; DB 12; Length 1275;
 Best Local Similarity 100.0%; Pred. No. 78;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GTGCAGCAGCTGGGCGCT 309
 |||||
 Db 463 GTGCAGCAGCTGGGCGCT 479

RESULT 39
 US-09-820-893-28/c
 : Sequence 28, Application US/09820893
 : Patent No. US20020076705A1
 : GENERAL INFORMATION:
 : APPLICANT: Rosen, et al.
 : TITLE OF INVENTION: 31 Human Secreted Proteins
 : FILE REFERENCE: P2033P1
 : CURRENT APPLICATION NUMBER: US/09/820,893
 : PRIOR FILING DATE: 2001-03-30
 : PRIOR APPLICATION NUMBER: 59/531,119
 : PRIOR FILING DATE: 2000-03-20
 : PRIOR APPLICATION NUMBER: 60/102,895
 : PRIOR FILING DATE: 1998-10-02
 : NUMBER OF SEQ ID NOS: 140
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 28
 : LENGTH: 1378
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 US-09-820-893-28

Query Match 1.1%; Score 17; DB 10; Length 1378;
 Best Local Similarity 100.0%; Pred. No. 79;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 964 CAAACCTCTGAGGAGC 980
 |||||
 Db 1417 CAAACCTCTGAGGAGC 1301

RESULT 40
 US-10-016-985-4
 : Sequence 3, Application US/10016985
 : Patent No. US20020123621A1
 : GENERAL INFORMATION:
 : APPLICANT: Walke, D. Wade
 : APPLICANT: Maricar, Miranda
 : APPLICANT: Yu, Xuanhuan (Sean)
 : APPLICANT: Fiddie, Carl Johan
 : TITLE OF INVENTION: No. US-985-123621A1 Human Kinase and Polynucleotides
 : TITLE OF INVENTION: Encoding the Same
 : FILE REFERENCE: LEX-0273-USA
 : CURRENT APPLICATION NUMBER: US/10/016,985
 : PRIOR FILING DATE: 2001-12-07
 : PRIOR APPLICATION NUMBER: US 60/251,941
 : NUMBER OF SEQ ID NOS: 3
 : SOFTWARE: FASTSEQ for Windows Version 4.0
 : SEQ ID NO 3
 : LENGTH: 1473
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 US-10-016-985-4

Query Match 1.1%; Score 17; DB 12; Length 1473;
 Best Local Similarity 100.0%; Pred. No. 79;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GTGCAGCAGCTGGGCGCT 309
 |||||
 Db 579 GTGCAGCAGCTGGGCGCT 595

RESULT 41
 US-09-808-602-26
 : Sequence 26, Application US/09808602
 : Patent No. US20020155115A1
 : GENERAL INFORMATION:
 : APPLICANT: Vornet, Corine A
 : APPLICANT: Fernandes, Elma
 : APPLICANT: Shimkets, Richard A
 : APPLICANT: Herrman, John L
 : APPLICANT: Majumder, Kamud
 : APPLICANT: Mishra, Vishnu
 : APPLICANT: Mezes, Peter S
 : APPLICANT: Macdonald, John
 : TITLE OF INVENTION: No. US20020155115A1 Proteins and Nucleic Acids Encoding Same
 : FILE REFERENCE: 15966-637 CIP
 : CURRENT APPLICATION NUMBER: US/09/808,602
 : PRIOR FILING DATE: 2001-03-14
 : PRIOR APPLICATION NUMBER: 69/800,198
 : PRIOR FILING DATE: 2001-03-05
 : PRIOR APPLICATION NUMBER: 50/180,590
 : FILE REFERENCE: LEX-0273-USA
 : CURRENT APPLICATION NUMBER: US/09/808,602
 : NUMBER OF SEQ ID NOS: 114
 : SOFTWARE: PatentIn Ver. 2.1
 : SEQ ID NO 26
 : LENGTH: 1498
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 US-09-808-602-26

Query Match 1.1%; Score 17; DB 9; Length 1498;
 Best Local Similarity 100.0%; Pred. No. 79;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 243 TGACATCTCCGCGAG 259
 |||||
 Db 1034 TGACATCTCCGCGAG 1050

RESULT 42
 US-09-848-726-1/c


```
Sequence 1, Application US/09848726
Patent No. US20020102667A1
GENERAL INFORMATION:
APPLICANT: NAMPOOHIRI, Madhavan
FILE OF INVENTION: No. US20020102667A1 Nucleotide Sequences Coding for the cbs Gen
CURRENT FILING DATE: 032301 WD 1191
CURRENT APPLICATION NUMBER: US/09/848,726
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1850
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
FEATURES:
NAME/KEY: CDS
LOCATION: (154)..(1653)
OTHER INFORMATION:
FEATURE:
NAME/KEY: RBS
LOCATION: (138)..(146)
OTHER INFORMATION:
US-09-848-726-1
```

```
Query Match
Best Local Similarity 100.0%; Score 17; DB 10; Length 1850;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1232 TTCAGACGCAATCTT 1248
Db 1281 TTCAGACGCAATAGTT 1295
```

```
RESULT 43
US-09-801-220-3
Sequence 3, Application US/09801220
Patent No. US20020173020A1
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel A.
APPLICANT: Williamson, Mark
TITLE OF INVENTION: 26866, A NOVEL CARBONITRILE ACYLTRANSFERASE
FILE REFERENCE: 10448-026001
CURRENT APPLICATION NUMBER: US/09/801,220
CURRENT FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 60/187,456
PRIOR FILING DATE: 2000-03-07
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 2412
TYPE: DNA
ORGANISM: Homo sapiens
US-09-801-220-3
```

```
Query Match
Best Local Similarity 100.0%; Score 17; DB 9; Length 2412;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 508 GCTGAGGAAATCTTACG 524
Db 81 GCTGAGGAAATCTTACG 97
```

```
RESULT 44
US-09-816-094-1
Sequence 1, Application US/09816094
Patent No. US20020064851A1
GENERAL INFORMATION:
APPLICANT: Wei, Ming-Hui et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
```

```
FILE REFERENCE: C1000536-CIP
CURRENT APPLICATION NUMBER: US20020161094
CURRENT FILING DATE: 2001-03-26
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2598
TYPE: DNA
ORGANISM: Human
US-09-816-094-1
```

```
Query Match
Best Local Similarity 100.0%; Score 17; DB 10; Length 2598;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 293 GTGACGACATGGGCT 309
Db 1236 GTGACGACATGGGCT 1252
```

```
RESULT 45
US-09-808-602-107
Sequence 107, Application US/09808602
Patent No. US20020155115A1
GENERAL INFORMATION:
APPLICANT: Verneil, Corine A
APPLICANT: Fernandes, Elma
APPLICANT: Shinkets, Richard A
APPLICANT: Herrman, John I.
APPLICANT: Majumder, Kunud
APPLICANT: Mishra, Vishnu
APPLICANT: Mezes, Peter S
APPLICANT: Macdougall, John
TITLE OF INVENTION: NO. US20020155115A1 Proteins and Nuc
FILE REFERENCE: 15966-697 CIP
CURRENT APPLICATION NUMBER: US200200960602
CURRENT FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 00/200,198
PRIOR FILING DATE: 2001-03-05
PRIOR APPLICATION NUMBER: 60/186,596
PRIOR FILING DATE: 2000-03-04
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 107
LENGTH: 2664
TYPE: DNA
ORGANISM: Homo sapiens
US-09-808-602-107
```

```
Query Match
Best Local Similarity 100.0%; Score 17; DB 9; Length 2664;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 243 TGAACATCTGGGCGAG 259
Db 672 TGAACATCTGGGCGAG 684
```

```
Search completed: December 7, 2002, 08:48:09
Job time : 569 secs
```

•
•
•
•

pred. No. is the number of results predicted by chance to have a

Result	No.	Score	Query Match	Length	bits	Info	File
C	1	556	36.1	1510	9	AE245219	AE
	2	518	34.5	1543	6	AX287118	AX
	3	518	33.6	1643	9	AE0015629	AE
	4	466	32.2	1800	9	AE290887	AE
	5	461	29.9	1200	9	AE042234	AX
	6	461	29.9	1535	9	AE042236	AX
	7	461	29.9	1548	9	AE042235	AX
	8	417	27.0	792	9	AE042240	AX
	9	417	27.0	999	9	AE042238	AX
	10	378	24.5	1472	9	AE042239	AX
C	11	378	24.5	4506	9	AE29048082	AE
	12	378	24.5	143619	9	AE008812	AX
	13	361	23.4	811	9	AE042237	AX
	14	345	23.4	190528	9	AE0424700	AX
	15	275	17.8	2921	9	AE0232750	AX
	16	220	14.3	506	6	AX048466	AX
	17	111	7.2	547	9	AE042233	AX
	18	111	7.2	867	9	AE042227	AX
	19	111	7.2	939	9	AE042223	AX
	20	111	7.2	994	9	AE042232	AX
C	21	111	7.2	1083	9	AE042226	AX
	22	111	7.2	1105	9	AE042241	AX
	23	111	7.2	1143	9	AE042225	AX
	24	111	7.2	1197	9	AE042222	AX
	25	111	7.2	1212	6	AE166105	AE
	26	111	7.2	1212	6	AE166106	AE
	27	111	7.2	1212	6	AX342903	AX
	28	111	7.2	1212	6	AX442604	AX
	29	111	7.2	1212	9	AE491085	AE
	30	111	7.2	1215	6	AX039403	AX
C	31	111	7.2	1215	9	AE042221	AX
	32	111	7.2	1243	9	AE042230	AX
	33	111	7.2	1212	6	AX287103	AX
	34	111	7.2	1312	9	HMLECTING	AX
	35	111	7.2	1315	9	AE042229	AX
	36	111	7.2	4566	9	AE290886	AE
	37	111	7.2	4424	9	AE209479	AE
	38	111	7.2	143619	9	AE290812	AE
	39	111	7.2	177962	9	AE008763	AE
	40	66	4.3	1146	9	AE469755	AE
C	41	63	4.3	1166	9	AE040419	AE
	42	63	4.3	1212	9	AE166107	AE
	43	63	4.3	1212	6	AX442905	AX
	44	59	3.8	1406	9	AE2948081	AE
	45	57	3.7	146	9	AE442727	AE

ALIT (5NM-F-N1.5)

RESULT 1	
AF245219	1610 bp mRNA linear
LOCUS	Homo sapiens probable mannose binding C-type lectin-like protein; complete cds.
DEFINITION	
ACCESSION	AF245219
VERSION	AF245219.2 GI:12084796
KEYWORDS	
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Euthera; Primates; Catarrhini; Hominoidea (Bases 1 to 1610)
AUTHORS	Sollieux,E.J., Barton,P. and Townsend,J.
TITLE	DCLIGN, a related gene, DCSIGN, and CD23 form a novel family of type I membrane proteins

DB	490	CCTGTGAAGGCTTCACTGGTTACTGTCGACATCGAATCCAAAGTGCGAGAGATCATACAC	549
OY	774	GAGCTGACCTGGGCTGAAGGCTGAGTAGTGGGTACTGATTGCACA	774
DB	550	GAGCTGACCGCCTCGAAGGCTGAGTGGGTGAATTGCCAGA	590
RESULT_6			
AV042236			
LOCUS	AV042236	1535 bp	mRNA
DEFINITION	Homo sapiens mDC-SIGN2 type IV isoform (CD209L) mRNA, complete cds.		
ACCESSION	AY042236		
VERSION	AY042236.1		
KEYWORDS	GI:15383609		
SOURCE			
ORGANISM	Homo sapiens.		
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 1535) Mummidl,S.; Ciano,G.; Lam,L.; Heelle,A.; Telles,V.; Ferguson,K.; Jimenez,F.; Anuja,S.S. and Ahuja,S.K. extensive repertoire of membrane-bound and soluble dendritic cell-specific ICAM-3-grabbing mannoprotein 1 (DC-SIGN) and DC-SIGN2 isoforms, inter-individual variation in expression of DC-SIGN transcripts		
AUTHORS	J. Biol. Chem. 276 (35), 34196-34212 (2001)		
TITLE			
JOURNAL	MEDLINE 21413847		
PUBMED	11337487		
REFERENCE	2 (bases 1 to 1535)		
AUTHORS	Mummidl,S. and Ahuja,S.K.		
TITLE	Direct Submission		
JOURNAL	Submitted (03-JUN-2001) Medicine, South Texas Veterans Health Care System and University of Texas Health Science Center at San Antonio, 7703 Floyd Curl Drive, San Antonio, TX 78229, USA		
FEATURES			
source	location/Qualifiers		
	1..1535		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="19"		
	/map="19p13.3"		
	1..1535		
	/gene="CD209L"		
	1..819		
	/gene="CD209L"		
	/note="Similar to the nucleotide sequence deposited in GenBank Accession Number AC008812; alternatively spliced"		
	/codon_start=1		
	/product="mDC-SIGN2 type IV isoform"		
	/protein_id="AAK91861.1"		
	/db_xref="GI:15383610"		
	/translation="MSDSKKERVOOLCHLEFDDPTSTACTPEPPDQPDQICHPSPSTC CLCHCALVQLSELPLAGVIVALLVGVNSVPSSLISGFQSDADAYONLTOLKAANG LSBKSLQEIYOELTOLKAANGELPEKSRLQEIYDELTRLKAAVELPEKSRLQEIYE LTRLKAAVGELPEKSRLQEIYDELTRLKAAVEGLPEKSRLQEIYDELTRLKAAVEGE PKSKLOEYIOELTOLKAANGELPDSKQQOQIYOELTLDTAFGRFLHKQPWA"		
gene			
CDS			
	1..819		
	/gene="CD209L"		
	1..819		
	/note="Similar to the nucleotide sequence deposited in GenBank Accession Number AC008812; alternatively spliced"		
	/codon_start=1		
	/product="mDC-SIGN2 type IV isoform"		
	/protein_id="AAK91861.1"		
	/db_xref="GI:15383610"		
	/translation="MSDSKKERVOOLCHLEFDDPTSTACTPEPPDQPDQICHPSPSTC CLCHCALVQLSELPLAGVIVALLVGVNSVPSSLISGFQSDADAYONLTOLKAANG LSBKSLQEIYOELTOLKAANGELPEKSRLQEIYDELTRLKAAVELPEKSRLQEIYE LTRLKAAVGELPEKSRLQEIYDELTRLKAAVEGLPEKSRLQEIYDELTRLKAAVEGE PKSKLOEYIOELTOLKAANGELPDSKQQOQIYOELTLDTAFGRFLHKQPWA"		
BASE COUNT	393 a 399 c 437 g 306 t		
QUERY MATCH	29.9%; Score 461; DB %; Length 1535;		
Best Local Similarity	100.0%; Pred. No. 9e+255;		
Matches 461; Conservative	0; Mismatches 0; Indels ; Gaps 0;		
OY	314	GCGGCTTGCGCATGGCGCCGCGGCTGCGCAATCCGTCTCTTCATGCGCTTGCG	373
DB	130	GCGGCTTTGGCATGATGGCGCCCTCGAGCTGTAACGTCGTCATGCGCTTGCG	189
OY	474	GTTCTGATGCTATCTTCTTCTTAAATGTCGAAGCTGCTTAACTCAAGAACAA	433
DB	190	GTTCTGATGCTATCTTCTTCTTAAATGTCGAAGCTGCTTAACTCAAGAACAA	249
OY	414	TTCGAGCAATGAGCGATCTATGCAATGCTATGATTAAATGTCGAAGCTGCT	493

DB	290	TTCAGGCGACAGCGAATCTATAAGAACTGACTGAGTTAAAGTGTCACTC
OY	494	TTCAGAGAATCCAAAGCTGTACAGACATTAACAGAGACCTGACCTGCAC
DB	310	TTCAGAGAAATCCAAAGCTGTACAGACATTAACAGAGACCTGACCTGCAC
OY	554	GGTGAGTGTCACAGAAATCCAAAGCTGTACAGAGATCTGCCAGAGCTGAC
DB	470	GGTGAGTGTCACAGAAATCCAAAGCTGTACAGAGATCTGCCAGAGCTGAC
OY	614	GCTGCATTGGTTCAGCTGTACAGAAATCCAAAGCTGTACAGAGATCTAAC
DB	430	GCTGCATTGGTTCAGCTGTACAGAAATCCAAAGCTGTACAGAGATCTAAC
OY	674	CGGTCGAAGCTGTGAGTGTGATGATGTCACAGAAATCCAAAGCTGTACAA
DB	490	CGGTCGAAGCTGTGAGTGTGATGATGTCACAGAAATCCAAAGCTGTACAA
OY	734	GAGCTCAACCTGGCTGGAAGGTCAGTACAGTAGTGTGATGTCACA
DB	550	GAGCTCAACCTGGCTGGAAGGTCAGTACAGTAGTGTGATGTCACA
RESULT 7		
AY042235		1648 bp mRNA linear
LOCUS		Homo sapiens mdc-slc2n2 type III isoform (CD209L)
DEFINITION		cdc, alternatively spliced.
ACCESSION		AY042235
VERSION		AY042235.1 GI:15484407
KEYWORDS		.
SOURCE		Homo sapiens.
ORGANISM		Homo sapiens.
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Homini
		1 (bases 1 to 1648)
		Mummidl,S., Catalao,R., Lam,L., Hoelle,A., Telles,
		Jimenez,F., Aluja,S.S., and Abuja,S.K.
		Extensive repertoire of membrane-bound and solubl
		cell-specific ICAM-3-grabing nonintegrin I (c ³)
		isoforms. Inter-individual variation in expressi
		transcripts
JOURNAL		J Biol Chem. 274 (45), 32396-32312 (2001)
MEDLINE		21413847
PUBMED		11337487
REFERENCE		2 (bases 1 to 1648)
AUTHORS		Mummidl,S. and Abuja,S.K.
TITLE		Direct Submission
JOURNAL		Submitted (03-JUN-2001) Medicine, South Texas V
		System and University of Texas Health Science Ce
		Antonio, 7703 Floyd Curl Drive, San Antonio, TX
		Location/Organism
FEATURES		
source		1..1648
		/organism "Homo sapiens"
		/db_xref "taxon:9606"
		/chromosome:"14"
		/map:"1p11.3"
		1..1648
gene		/gene="CD209L"
CDS		1..819
		/gene="CD209L"
		/note="similar to the nuc-polide sequen
		GenBank Accession Number AF008812; att
		/codon_start=1
		/product "mdc-slc2n2 type III isoform"
		/protein_id="AAK91860.1"
		/db_xref "GI:15483608"
		/translation "MSEKTEFEWQDILDEETSGH
		LGEHAIVGHISMTAGVVAIVAVSVPSLSCH
		LSERKLDTIYRIITKAANGELPKRSKIPTIGTIR
		ELTRKAANGELPKRSKIPTIGTIRLKAANGELPFRS-
		PKRSKIPTIGTIRLKAANGELPDGSKDOOIVDFITR-
BASE COUNT		420 a 432 c 446 g 430 t

exon		/rpt_type=tandem /rpt_unit=613..681 1557..1708 /gene="abc" note="encodes c-type lectin domain CRD involved in the formation of the carbohydrate recognition domain"
exon		/number=5 2117..2529 /gene="abc" note="encodes c-type lectin domain CRD involved in the formation of the carbohydrate recognition domain"
exon		/number=6 3740..3945 /gene="abc" note="encodes c-type lectin domain CRD involved in the formation of the carbohydrate recognition domain"
exon		/number=7 4275..44506 /gene="abc" note="encodes c-type lectin domain CRD involved in the formation of the carbohydrate recognition domain"
BASE COUNT	1106 a	1091 c 1291 g 1017 t 1 others
ORIGIN		
Query Match	24.5%; Score 378; DB 9; Length 4506;	
Best Local Similarity	100.0%; Pred No 1a-206;	
Matches 378, Conservative 0, Mismatches 0; Indels 0; Gaps 0;		
OY	397 AGGTGTAAGATGGTCTACGTTTCCTAAAGTCAGAACCAATCCAGACAGCACAATCTACCA 456	
Dh	537 AGTGTGAAGAGTGCGGAGCTGCGTAAAGTCAGAACCAATCTACAGAAATATCATCTACCA 596	
OY	457 GAACCTCAACCTCAACTTAAAGTCTGAGTGCATCATCTTCAATAAATCTAAATCTACCA 516	
Ld	597 GAACCTCAACCTCAACTTAAAGTCTGAGTGCATCATCTTCAATAAATCTAAATCTACCA 656	
OY	517 GATGTATCACAGCTGACGACCATCGAAGAGCTCTGAGTGCATCATCTTCAATAAATCTACCA 576	
Dh	657 GATGTATCACAGAGTGTACCTAGATCAAGATCTGAGTGCATCATCTTCAATAAATCTACCA 716	
OY	577 GTGTATCAAT 636	
Dh	717 GTGTATCAAT 776	
OY	637 GAAATGTAATATCTGTAT 696	
Ld	777 TAAATGTAATATCTGTAT 836	
OY	697 ATTCTACATTAAT 756	
Dh	837 ATCTACATTAAT 896	
OY	757 AGTGGGTGAGTGTGACCA 774	
Dh	897 AGTGGGTGAGTGTGACCA 914	
RESULT 12		
AC008812	147619 bp DNA linear FR: 28 JUL 2000	
LOCUS	AC008812 Homo sapiens chromosome 19 clone CTD-210c119, complete sequence.	
DEFINITION	Homo sapiens chromosome 19 clone CTD-210c119, complete sequence.	
ACCESSION	AC008812.7 GI:9558577	
VERSION	AC008812.7 GI:9558577	
KEYWORDS	HTG.	
SOURCE	Homo sapiens.	
ORGANISM	Homo sapiens.	
REFERENCE	Eukariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia, Eutheria, Primates, Catarrhini, Homiidae, Homo. 1 (bases 1 to 143619) Lab Joint Genome Institute and Stanford Human Genome Center. Direct Submission Unpublished 2 (bases 1 to 143619) DOE Joint Genome Institute.	
AUTHORS		

TITLE	Direct Submission
JOURNAL	Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE	3 (Bases 1 to 147619)
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE	Direct Submission
JOURNAL	Submitted (28-JUL-2000) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT	on Jul 28, 2000 this sequence version replaced gi.7711353. Draft Sequence Produced by DOE Joint Genome Institute www.jgi.doe.gov Finishing Completed at Stanford Human Genome Center www.shgc.stanford.edu Quality: Phrap Quality >=40 99.9% of Sequence; Estimated Total Number of Errors is 3.8.
FEATURES	
SOURCE	Location/Qualifiers L_1143619 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="19" /clone="CTD-2102P19"
BASE COUNT	36628 a 35385 c 35414 g 36192 t
ORIGIN	
Query Match	24.5%; Score 378; DB %; Length 143619;
Best Local Similarity	100 0%; Proq No q 4e-207;
Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0.	
OY 347	AGTGTCAAGGTCCTCCAGCTCCTTAAGTCACAAATCGACGACACGCCAATCTACCA 456 Db 68747 AGTGCCAGAGTCCCGACGCTCCTTAAGTAGAGAACAAATCCGACGAAGAAGCTAACTACCA 68806
OY 457	GAACTGCACCACACTTAAGCTTCGACGTGGGTACCTCTTCAGAGAAATCCAGCTGCAGCA 516 Db 68807 GAACCTTACCTCAATTAAAGCTCTCAAGTGGAGCTCAGCTCAGCAAAATCCAGCTGCAGCA 68866
OY 517	GATCTACCAAGAGCTGACCTGACCTTAAGGCTGCATGATGATAGTTTCGACAGAGAAATCCAA 576 Db 68867 GATTTACCTAGGAGCTTACCTGACCTTAAGGCTGCATGATGATAGTTTCGACAGAGAAATCCAA 68926
OY 577	GCTCCAGAGAGATTAATGAGAGGCTGACCTGCTGTAAGCTGCTGAGTCTAGTTCGATGCACA 636 Db 68927 GTCTAGGAGATTTAATAAGCTGCTGTAAGCTGCTGAGTCTAGTTCGATGCACA 68986
OY 647	GAATTCGACAGCTGACGAGATCTTACGACAGAGCTGACCTGCTGTAAGCTGCTGAGTTCGATGC 696 Db 68947 GAATTCGACAGCTGACGAGATCTTACGACAGAGCTGACCTGCTGTAAGCTGCTGAGTTCGATGC 69046
OY 657	GTTCCTAGAGAAAATCCAGAGTGGAGATATTAATGAGAGCTGAGTGGCTGAGAGGTTGGT 756 Db 69047 GTTCCTAGAGAAAATCCAGAGTGGAGATATTAATGAGAGCTGAGTGGCTGAGAGGTTGGT 69106
OY 757	AGTGTGTGAGTTCGCAGCA 774
Db 69107	AGTGTGTGAGTTCGCAGCA 69124
RESULT 13	
AY042237	811 bp mRNA linear PKI 30-AUG-2001
DEFINITION	Homo sapiens mDC-SIGN2 type VI isoform (CD299L) mRNA, complete cds, alternatively spliced.
ACCESSION	AY042237
VERSION	AY042237.1 GI:15483611
KEYWORDS	.
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	Mummidhi,S., Catanoso,G., Lam,L., Hoeft,A., Torres,V., Ferguson,K., Jimenez,F., Anuja,S.S. and Anuja,S.K.
TITLE	Extensive repertoire of membrane-toumd and soluble dendritic

FEATURES	source
JOURNAL	cell-specific iCAM-Targeting nonintegrin 1 (IC-1)
MEDLINE	transcripts
PUBMED	J. Biol. Chem. 276 (45), 42196-42212 (2001)
REFERENCE	21413847
AUTHORS	11337487
TITLE	2 (bases 1 to 811)
JOURNAL	Mumidi, S. and Ahuja, S. K.
	Direct Submission
	Submitted (03-JUN-2001) Medicine, South Texas Vet.
	System and University of Texas Health Science Cen-
	Antonio, 7703 Floyd Curl Drive, San Antonio, TX 782
	Location/Qualifiers
	1..811
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="19"
	/map="19p13.3"
	1..811
	/gene="CD209L"
	1..699
	/gene="CD209L"
	/note="Similar to the nucleotide sequence
	genbank Accession Number AF008812; 327402
	/codon_start 1
	/product="cell surface type VI isoform"
	/protein="14 "AAK3462.1"
	/db_xref="gi:1548362.1"
	/translation="MSVSKRPVQDLEEDPIISIRK
	CLGRLAVLQISPMILAGVAVLAVLAVSKPSSLSQSS
	LSFSSKLQETVDELIGKAAVGLPEKSKLQETVDELIGL
	ELTDKAAVERIYPRHPRITWTFQANVFVMSNSQKPMH
	SLFQLVAVLEQWPIQQ"
BASE COUNT	213 a 218 c 226 g 154 t
ORIGIN	
Query Match	23.4%; Score 361; DB 9; Length 8
Best Local Similarity	100.0%; Pred. No. 7,44-197;
Matches 361; Conservative 9; Mismatches 0; Indels	
314 GAGGCTCTTGAGCAATGGCGCGCTGGGCTGAACTCTGCTTCATGACGTC	
130 GATGATCTTTGACATAGCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	
374 GTGCTGCTGGACGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	
130 GTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	
434 TCCGAGCAAGACGCAATCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	
230 TCCGAGCAAGACGCAATCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	
494 TCAGAGAAATCCAAAGCTGATAGAGAGATAGAGAGAGAGAGAGAGAGAGAGAG	
310 TCAGAGAAATCCAAAGCTGATAGAGAGATAGAGAGAGAGAGAGAGAGAGAGAG	
534 GTTGAGTTGGCAGAGAAATCCAAAGCTGATAGAGAGATAGAGAGAGAGAGAG	
370 GTTGAGTTGGCAGAGAAATCCAAAGCTGATAGAGAGATAGAGAGAGAGAGAG	
614 GCTGAGTATGAGAGATCTGATAGAGAGATAGAGAGATAGAGAGATAGAGAG	
430 GCTGAGTATGAGAGATCTGATAGAGAGATAGAGAGATAGAGAGATAGAGAG	
674 C 674	
490 C 490	

SEQUENCE, 32 unordered pieces.

AC024700
 VERSION
 AC024700.4 31:856979
 HITS: HITS_PHASE1: HITS_PHASE1
 KIWIKIDS
 SOURCE
 Homo sapiens.
 ORGANISM
 Homo sapiens.
 Mammalia: Eutheria: Chordata: Craniata: Vertebrata: Eumetazoa: Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
 1 (bases 1 to 190628)
 Waterston, R.H.
 The sequence of Homo sapiens clone
 unpublished
 2 (bases 1 to 190629)
 Waterston, R.H.
 Direct Submission
 Submitted (01-MAR-2000) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, SE, Louis-
 Mo 63108, USA
 On Jan 16, 2000 this sequence version replaced gi:7523915.

Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc/index.shtml>
 Project information
 Center project name: H.NH0042118
 Summary Statistics
 Sequencing vector: M13: 100%

Sequencing vector: plasmid: 0%

Chemistry: Dye-primer ET, 100% of reads

Chemistry: Dye-terminator Big Dye, 0% of reads

Assembly program: Phrap: version 0.990319

Consensus quality: 173687 bases at least 930

Consensus quality: 181172 bases at least 920

Consensus quality: 181172 bases at least 920

Insert size: 175000; agarose-gel

Insert size: 187528; sum-of-contigs

Quality coverage: 5.24 in 920 bases; sum-of-contigs

Quality coverage: 5.00 in 920 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently
 consists of 32 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available, and the accession number will
 be preserved.

1 1244: contig of 1244 bp in length
 1245 1344: gap of unknown length
 1345 2396: contig of 1052 bp in length
 2397 2406: gap of unknown length
 2497 3585: contig of 1089 bp in length
 3586 3685: gap of unknown length
 3686 5158: contig of 1473 bp in length
 5159 5298: gap of unknown length
 5299 6625: contig of 1326 bp in length
 6626 6724: gap of unknown length
 6725 8401: contig of 1677 bp in length
 8402 8501: gap of unknown length
 8502 10172: contig of 1671 bp in length
 10173 10272: gap of unknown length
 10273 12510: contig of 2238 bp in length
 12511 12610: gap of unknown length
 12611 14735: contig of 2125 bp in length
 14736 14835: gap of unknown length
 14836 16444: contig of 1608 bp in length
 16444 16543: gap of unknown length
 16544 19269: contig of 2726 bp in length
 19270 19369: gap of unknown length
 19370 21690: contig of 2321 bp in length
 21691 21790: gap of unknown length
 21791 25543: contig of 3753 bp in length

25544 25643: gap of unknown length
 25644 28328: contig of 2885 bp in length
 28329 28428: gap of unknown length
 28429 31639: contig of 3311 bp in length
 31640 31739: gap of unknown length
 31740 34456: contig of 4727 bp in length
 34457 36566: gap of unknown length
 36567 39593: contig of 3027 bp in length
 39594 39693: gap of unknown length
 39694 46098: contig of 6405 bp in length
 46099 46198: gap of unknown length
 46199 50494: contig of 4296 bp in length
 50495 50595: gap of unknown length
 50596 56467: contig of 5773 bp in length
 56468 56567: gap of unknown length
 56568 61903: contig of 5436 bp in length
 61904 62003: gap of unknown length
 62004 67159: contig of 5156 bp in length
 67160 67260: gap of unknown length
 67261 74171: contig of 6912 bp in length
 74172 74271: gap of unknown length
 74272 80579: contig of 6308 bp in length
 80580 80679: gap of unknown length
 80680 88540: contig of 7861 bp in length
 88541 88640: gap of unknown length
 88641 95519: contig of 6878 bp in length
 95520 95619: gap of unknown length
 95620 102484: contig of 6866 bp in length
 102485 102584: gap of unknown length
 102585 109466: contig of 6884 bp in length
 109467 109566: gap of unknown length
 109567 130223: contig of 20654 bp in length
 130224 130323: gap of unknown length
 130324 154893: contig of 24567 bp in length
 154894 154993: gap of unknown length
 154994 171709: contig of 16720 bp in length
 171710 171809: gap of unknown length
 171810 171811: contig of 18819 bp in length.

FEATURES
 SOURCE
 1..190628
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="19"
 /colore="P11-42018"
 1..1244
 /note="assembly_name:Contig15"
 1245..2396
 /note="assembly_name:Contig16"
 2397..3585
 /note="assembly_name:Contig17"
 3586..5158
 /note="assembly_name:Contig20"
 5159..6625
 /note="assembly_name:Contig23"
 6626..8401
 /note="assembly_name:Contig24"
 8402..10172
 /note="assembly_name:Contig25"
 10173..12510
 /note="assembly_name:Contig29"
 12511..14735
 /note="assembly_name:Contig30"
 14736..16444
 /note="assembly_name:Contig31"
 16445..19269
 /note="assembly_name:Contig32"
 19270..21690
 /note="assembly_name:Contig33"
 21691..25543
 /note="assembly_name:Contig34"
 25544..3753
 /note="assembly_name:Contig35"
 3754..458a-25.01.rge


```

misc_feature 1..126
/gene="CD209"
/note="putative open reading frame, encodes truncated
protein TDC-SIGN1b"
/evidence not-experimental
101..994
/gene="CD289"
/note="similar to the nucleotide sequence deposited in
GenBank Accession Number AC008812; alternatively spliced"
/codon_start=1
/product="sbc-sig1nb type III isoform"
/protein_id="AAK91857.1"
/db_xref="GI:15281095"
/translat=1..4..*MASA-P-SSTFTSIHSEFEGLPLDPEGLPEVYSIAVSFVSISL
OOSRODIAIYONITOLKAAGVGLPEKSKLOIYQELITLKAAGVGLPEKSKLOIYQEL
LITLKAAGVGLPEKSKLOIYQELITLKAAGVGLPEKSKLOIYQELITLKAAGVGL
EKSROOITLITLKAAGVGLPEKSKLOIYQELITLKAAGVGLPEKSKLOIYQEL
TOLKAVERI/HPCEWETFEQNTYFMSNSQPMWHDSTICEVQACIIVIKSAFEQ
LQAVLEORPAQWNRGRLPPI"
BASE COUNT 265 a 244 c 310 g 175 t
ORIGIN

Query Match 7.28; Score 111; DB 9; Length 994;
Best Local Similarity 100.0%; Pred. No. 4.4e-52;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q7 449 ATCTACCAAGAACCTGACCTGAGTTAAAGCTGAGTGGTGATCTTCAGCAAAATCGAAG 508
|||||
D3 257 ATCTACCAAGAACCTGACCTGAGTTAAAGCTGAGTGGTGATCTTCAGCAAAATCGAAG 316
|||||

Q7 509 CTGAGAGAGATCTACCAAGAGCTGAGTAAAGCTGAGTGGTGATCTTCAGCAAAATCGAAG 559
|||||
D3 317 CTGAGAGAGATCTACCAAGAGCTGAGTAAAGCTGAGTGGTGATCTTCAGCAAAATCGAAG 367
|||||

RESULT 21
AY042226 1083 bp mRNA linear PRI 28 AUG 2001
LOCUS Homo sapiens SDC-SIGN1A type II isoform (CD209) mRNA, complete cds,
alternatively spliced
ACCESSION AY042226
VERSION AV042226.1 GI:15281082
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 1083)
Mummidli,S., Catano,G., Lam,L., Hoeft,A., Teiles,V., Begum,K.,
Jimenez,F., Ahuja,S.S., and Ahuja,S.K.
Extensive repertoire of membrane-bound and soluble dendritic
cell-specific ICAM-3-grabbing nonintegrin 1 (DC-SIGN) and DC-SIGN2
isoforms, inter-individual variation in expression of DC-SIGN
transcripts
J Biol Chem. 276 (45): 43196-43212 (2001)

JOURNAL 21413847
MEDLINE 1137487
PUBMED 2 (bases 1 to 1083)
REFERENCE Mummidli,S. and Ahuja,S.K.
AUTHORS Direct Submission
TITLE Submitted (18-JUN-2001) Department of Medicine, South Texas
Veterans Health Care System and University of Texas Health Science
Center, 7703, Floyd Curl Drive, San Antonio, TX 78229, USA
JOURNAL Location/Qualifiers
FEATURES
SOURCE 1..1083
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19p13.3"
1..1083
/gene="CD209"
1..1083
/gene="CD209"
/note="similar to the nucleotide sequence deposited in
GenBank Accession Number AC008812; alternatively spliced"
/codon_start=1
/product="sbc-sig1nb type II isoform"
/protein_id="AAK91851.1"
/db_xref="GI:15281083"
/translat=1..4..*MSGKEPPI20V/ILTSKVSRSISGEFSDADAIYONITOLKAAG
VGLSEKSKLOIYQELITLKAAGVGLPEKSKLOIYQELITLKAAGVGLPEKSKLOIY
VGLITLKAAGVGLPEKSKLOIYQELITLKAAGVGLPEKSKLOIYQELITLKAAG
ELPEKSKLOIYQELITLKAAGVGLPEKSKLOIYQELITLKAAGVGLPEKSKLOIY
FOGNYFMSNSQPMWHDSTICEVQACIIVIKSAFQNP/IOSSSNFTWWSID
LNOEGTWQWVDGSDPLLPSEKQYMNRGEPNNVEEDCAEFSNGMNIDKMLIAFWICK
KSAASKEEFILSPAPATPMPPA"
BASE COUNT 290 a 284 c 310 g 199 t
ORIGIN

Query Match 7.28; Score 111; DB 9; Length 1083;
Best Local Similarity 100.0%; Pred. No. 4.4e-52;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q7 449 ATCTACCAAGAACCTGACCTGAGTTAAAGCTGAGTGGTGATCTTCAGCAAAATCGAAG 508
|||||
D3 97 ATCTACCAAGAACCTGACCTGAGTTAAAGCTGAGTGGTGATCTTCAGCAAAATCGAAG 156
|||||

Q7 509 CTGAGAGAGATCTACCAAGAGCTGAGTAAAGCTGAGTGGTGATCTTCAGCAAAATCGAAG 559
|||||
D3 157 CTGAGAGAGATCTACCAAGAGCTGAGTAAAGCTGAGTGGTGATCTTCAGCAAAATCGAAG 207
|||||

RESULT 22
AY042231 1105 bp mRNA linear PRI 28 AUG 2001
LOCUS Homo sapiens SDC-SIGN1B type II isoform (CD209) mRNA, complete cds,
alternatively spliced
ACCESSION AY042231
VERSION AV042231.1 GI:15281092
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 1105)
Mummidli,S., Catano,G., Lam,L., Hoeft,A., Teiles,V., Begum,K.,
Jimenez,F., Ahuja,S.S., and Ahuja,S.K.
Extensive repertoire of membrane-bound and soluble dendritic
cell-specific ICAM-3-grabbing nonintegrin 1 (DC-SIGN) and DC-SIGN2
isoforms, inter-individual variation in expression of DC-SIGN
transcripts
J Biol Chem. 276 (45): 43196-43212 (2001)

JOURNAL 21413847
MEDLINE 1137487
PUBMED 2 (bases 1 to 1105)
REFERENCE Mummidli,S. and Ahuja,S.K.
AUTHORS Direct Submission
TITLE Submitted (18-JUN-2001) Department of Medicine, South Texas
Veterans Health Care System and University of Texas Health Science
Center, 7703, Floyd Curl Drive, San Antonio, TX 78229, USA
JOURNAL Location/Qualifiers
FEATURES
SOURCE 1..1105
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19p13.3"
1..1105
/gene="CD209"
1..126
/gene="CD209"
/note="putative open reading frame, encodes truncated
protein TDC-SIGN1b"
/evidence not-experimental
101..1105
/gene="CD209"
/note="similar to the nucleotide sequence deposited in
GenBank Accession Number AC008812; alternatively spliced"
/codon_start=1
/product="sbc-sig1nb type II isoform"
/protein_id="AAK91851.1"
/db_xref="GI:15281083"
/translat=1..4..*MSGKEPPI20V/ILTSKVSRSISGEFSDADAIYONITOLKAAG
VGLSEKSKLOIYQELITLKAAGVGLPEKSKLOIYQELITLKAAGVGLPEKSKLOIY
VGLITLKAAGVGLPEKSKLOIYQELITLKAAGVGLPEKSKLOIYQELITLKAAG
ELPEKSKLOIYQELITLKAAGVGLPEKSKLOIYQELITLKAAGVGLPEKSKLOIY
FOGNYFMSNSQPMWHDSTICEVQACIIVIKSAFQNP/IOSSSNFTWWSID
LNOEGTWQWVDGSDPLLPSEKQYMNRGEPNNVEEDCAEFSNGMNIDKMLIAFWICK
KSAASKEEFILSPAPATPMPPA"
BASE COUNT 290 a 284 c 310 g 199 t
ORIGIN

```


GenBank Accession Number AC098812, alternatively spliced"

/codon_start=1
/product="SDC-SIGNIA type II isoform"
/protein_id="AAK91856.1"
/db_xref="GI:15281093"

BASE COUNT 288 a 289 c 321 g 207 t

Query Match 7.2% Score 111; Db % Length 1105;
Best local Similarity 100.0%; Pred. No. 4,4e-52;
Matches 111: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 449 ATCTACGAGAGCTGACCTTAAGCTTGAAGTGGAGTGTCTGAGCAATGCAAG 508
|||||
DB 257 ATCTACGAGAGCTGACCTTAAGCTTGAAGTGGAGTGTCTGAGCAATGCAAG 316
|||||

QY 509 CTGACGAGATCTACGAGATCTACGAGATCTACGAGATCTACGAGATCTACGAG 556
|||||
DB 417 CTGACGAGATCTACGAGATCTACGAGATCTACGAGATCTACGAGATCTACGAG 457
|||||

RESULT 23
AY042225 1143 bp mRNA linear Feb 28/00/2001
LOCUS Homo sapiens SDC-SIGNIA type I isoform (CD209) mRNA, complete cds,
alternatively spliced.
ACCESSION AY042225
VERSION AY042225.1 GI:15281080
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo:
1 (bases 1 to 1143)
Mummidi, S., Catano, G., Lam, L., Hoelle, A., Telles, V., Begum, K.,
Jimenez, F., Ahuja, S.S., and Ahuja, S.K.
Extensive repertoire of membrane bound and soluble dendritic
cell-specific ICAM-3-grabbing nonintegrin 1 (DC-SIGN) and DC-SIGN2
isoforms. Inter-individual variation in expression of DC-SIGN
transcripts
JOURNAL 1. Biol. Chem. 276 (35): 33196-33212 (2001)
MEDLINE 21413847
PubMed 11337487
REFERENCE 2 (bases 1 to 1143)
Mummidi, S., and Ahuja, S.K.
Direct Submission
TITLE Submitted (18-JUN-2001) Department of Medicine, South Texas
Veterans Health Care System and University of Texas Health Science
Center, 7703, Floyd Curl Drive, San Antonio, TX 78229, USA
AUTHORS
JOURNAL
FEATURES
SOURCE
1. 1143
/organism "Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19p13.3"
1. 1143
/gene="CD209"
1. 1143
/gene="CD209"
/note="similar to the nucleotide sequence deposited in
GenBank Accession Number AC098812; alternatively spliced"
/codon_start=1
/product="SDC-SIGNIA type I isoform"
/protein_id="AAK91850.1"
/db_xref="GI:15281081"
/translation="MSDSKEPRLLQGLIFEPFLGLGFLFGLFGLYGLKSLAVSKVPSSTIS

QEOSKQDAIVNLTOLKAAGVSEKSKLQELIYGLTQETL
LIRLKAAGVLEPKSKLQELIYGLTQELIYGLTQELIYGLTQELIYGLTQELIYGLTQETL
EKSKQELIYGLTQELIYGLTQELIYGLTQELIYGLTQELIYGLTQELIYGLTQELIYGLTQETL
TOLKAAGVLEPKSKLQELIYGLTQELIYGLTQELIYGLTQELIYGLTQELIYGLTQETL
NLOLOSRSRSHFTWMTFPGNCFKNSORNMHDSITL
PSNQMNIYCNIAKEWTKKSAASCPDEDEFLSPAPA
BASE COUNT 409 a 296 c 330 g 208 t

Query Match 7.2% Score 111; Db % Length 1
Best local Similarity 100.0%; Pred. No. 4,4e-52;
Matches 111: Conservative 0; Mismatches 0; Indels 0;

QY 449 ATCTACGAGAGCTGACCTTAAGCTTGAAGTGGAGTGTCTGAGCAATGCAAG 508
|||||
DB 217 CTGACGAGATCTACGAGATCTACGAGATCTACGAGATCTACGAGATCTACGAG 457
|||||

RESULT 24
AY042222 1197 bp mRNA linear Feb 28/00/2001
LOCUS Homo sapiens MDC-SIGNIA type II isoform (CD209) mRNA,
alternatively spliced.
ACCESSION AY042222
VERSION AY042222.1 GI:15281074
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo:
1 (bases 1 to 1197)
Mummidi, S., Catano, G., Lam, L., Hoelle, A., Telles, V., Begum, K.,
Jimenez, F., Ahuja, S.S., and Ahuja, S.K.
Extensive repertoire of membrane bound and soluble
cell-specific ICAM-3-grabbing nonintegrin 1 (DC-SIGN) and DC-SIGN2
isoforms. Inter-individual variation in expression of DC-SIGN
transcripts
JOURNAL 1. Biol. Chem. 276 (35): 33196-33212 (2001)
MEDLINE 21413847
PubMed 11337487
REFERENCE 2 (bases 1 to 1197)
Mummidi, S., and Ahuja, S.K.
Direct Submission
TITLE Submitted (18-JUN-2001) Department of Medicine, South Texas
Veterans Health Care System and University of Texas Health Science
Center, 7703, Floyd Curl Drive, San Antonio, TX 78229, USA
AUTHORS
JOURNAL
FEATURES
SOURCE
1. 1197
/organism "Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19p13.3"
1. 1197
/gene="CD209"
1. 1197
/gene="CD209"
/note="similar to the nucleotide sequence
GenBank Accession Number AC098812; alter

BASE COUNT 310 a 315 c 346 g 226 t
ORIGIN

Query Match 7.2%; Score 111; DB 9; Length 1197;
Best Local Similarity 100.0%; Pred. No. 4.4e-52;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 449 ATCTACCAACCTGACCTTAAAGCTGAGGAGGACCTGTGAGAAATCCAAAG 508
|||||
DB 229 ATCTACCAACCTGACCTTAAAGCTGAGGAGGACCTGTGAGAAATCCAAAG 288

QY 509 CTGTACCAACCTGACCTTAAAGCTGAGGAGGACCTGTGAGAAATCCAAAG 559
|||||
DB 289 CTGTACCAACCTGACCTTAAAGCTGAGGAGGACCTGTGAGAAATCCAAAG 349

RESULT 25
AR166105 1212 bp DNA linear PAT 17-OCT-2001
LOCUS
DEFINITION Sequence 9 from patent US 6280953.
ACCESSION AP166105
VERSION AP166105.1 GI:16241278
KEYWORDS
ORGANISM Unknown.
SOURCE Unclassified.
REFERENCE 1 (bases 1 to 1212)
AUTHORS Messier W. and Sikela J.M.
TITLE Methods to identify polynucleotide and polypeptide sequences which may be associated with physiological and medical conditions
JOURNAL Patent: US 6280953-A 9 28-AUG-2001.
FEATURES
Source 1..1212
Location/Qualifiers
/organism="unknown"

BASE COUNT 314 a 321 c 348 g 229 t
ORIGIN

Query Match 7.2%; Score 111; DB 6; Length 1212;
Best Local Similarity 100.0%; Pred. No. 4.4e-52;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 449 ATCTACCAACCTGACCTTAAAGCTGAGGAGGACCTGTGAGAAATCCAAAG 508
|||||
DB 229 ATCTACCAACCTGACCTTAAAGCTGAGGAGGACCTGTGAGAAATCCAAAG 288

QY 509 CTGTACCAACCTGACCTTAAAGCTGAGGAGGACCTGTGAGAAATCCAAAG 559
|||||
DB 289 CTGTACCAACCTGACCTTAAAGCTGAGGAGGACCTGTGAGAAATCCAAAG 349

RESULT 26
AR166106 1212 bp DNA linear PAT 17-OCT-2001
LOCUS
DEFINITION Sequence 10 from patent US 6280953.
ACCESSION AR166106
VERSION AR166106.1 GI:16241278
KEYWORDS
ORGANISM Unknown.
SOURCE Unclassified.
REFERENCE 1 (bases 1 to 1212)
AUTHORS Messier W. and Sikela J.M.
TITLE Methods to identify polynucleotide and polypeptide sequences which may be associated with physiological and medical conditions
JOURNAL Patent: US 6280953-A 10 28-AMT-2001;
FEATURES
Source 1..1212
Location/Qualifiers
/organism="unknown"

BASE COUNT 315 a 316 c 350 g 230 t 1 others
ORIGIN

Query Match 7.2%; Score 111; DB 6; Length 1212;

Best Local Similarity 100.0%; Pred. No. 4.4e-52;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 449 ATTACCAACCTGACCTTAAAGCTGAGGAGGACCTGTGAGAAATCCAAAG 508
|||||
DB 229 ATTACCAACCTGACCTTAAAGCTGAGGAGGACCTGTGAGAAATCCAAAG 288

QY 509 CTGTACCAACCTGACCTTAAAGCTGAGGAGGACCTGTGAGAAATCCAAAG 559
|||||
DB 289 CTGTACCAACCTGACCTTAAAGCTGAGGAGGACCTGTGAGAAATCCAAAG 349

RESULT 27
AX342003 1212 bp DNA linear PAT 10-JAN-2002
LOCUS
DEFINITION Sequence 9 from Patent W00196603.
ACCESSION AX342003
VERSION AX342003.1 GI:18137981
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 1212)
AUTHORS Messier W. and Sikela J.M.
TITLE Methods to identify polynucleotide and polypeptide sequences which may be associated with physiological and medical conditions
JOURNAL Patent: W0196603-A 9 20-JAN-2001;
FEATURES
Source 1..1212
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 314 a 321 c 348 g 229 t
ORIGIN

Query Match 7.2%; Score 111; DB 6; Length 1212;
Best Local Similarity 100.0%; Pred. No. 4.4e-52;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 449 ATTACCAACCTGACCTTAAAGCTGAGGAGGACCTGTGAGAAATCCAAAG 508
|||||
DB 229 ATTACCAACCTGACCTTAAAGCTGAGGAGGACCTGTGAGAAATCCAAAG 288

QY 509 CTGTACCAACCTGACCTTAAAGCTGAGGAGGACCTGTGAGAAATCCAAAG 559
|||||
DB 289 CTGTACCAACCTGACCTTAAAGCTGAGGAGGACCTGTGAGAAATCCAAAG 349

RESULT 28
AX342004 1212 bp DNA linear PAT 10-JAN-2002
LOCUS
DEFINITION Sequence 10 from Patent W00196603.
ACCESSION AX342004
VERSION AX342004.1 GI:18137982
KEYWORDS
SOURCE chimpanzee.
ORGANISM Pan troglodytes.
REFERENCE 1 (bases 1 to 1212)
AUTHORS Messier W. and Sikela J.M.
TITLE Methods to identify polynucleotide and polypeptide sequences which may be associated with physiological and medical conditions
JOURNAL Patent: W0196603-A 10 20-DEC-2001;
FEATURES
Source 1..1212
Location/Qualifiers
/organism="Pan troglodytes"
/db_xref="taxon:9598"

BASE COUNT 315 a 316 c 350 g 230 t 1 others
ORIGIN

[illegible][illegible]

LOCUS	AX287103	1122 bp	DNA	linear	FAT 21 NOV 2001
DEFINITION	Sequence 1 from Fat2d Web04702.				
ACCESSION	AX287104				
VERSION	AX287104.1	GI:17045083			
KEYWORDS					
SOURCE	Human.				
ORGANISM	Homo sapiens				
REFERENCE	Elkayam, M., Yaron, O., Chertkov, G., Golan, Y., Verheggen, J., Entchev, S., Mammalia, E., Eshkol, E., Finkels, G., Golan, Y., and Gelfand, E. (2001). Antibody inhibiting the binding between p120 and cyclin and screening methods.				
AUTHORS	Elkayam, M., Yaron, O., Chertkov, G., Golan, Y., Verheggen, J., Entchev, S., Mammalia, E., Eshkol, E., Finkels, G., Golan, Y., and Gelfand, E.				
TITLE	Antibody inhibiting the binding between p120 and cyclin and screening methods.				
JOURNAL	Patent WO 0164752-A 1 (7-SEP-2001); New York University (US); KATHOLIEKE UNIVERSITEIT NIJMEGEN (NL)				
FEATURES					
SOURCE	location/Qualifiers				
1..1312					
/organism="Homo sapiens"					
/db_xref="GI:17045083"					
42..1256					
/note="unnamed protein product"					
/codon_start=1					
/protein_id="CAI2297.1"					
/db_xref="GI:17045084"					
/translation="MDSKREKPPDGGGLPEPQVGPQPPQYKSLACLGAGPVL					
EQLQSTLTLALVQWSKREKPSISQEDSPQADALVTLKAAVGLSEPKSLQEIYQ					
ELTQKAAVGLPEKSKLQEIYQHLTRKAAVGLPEKSKLQEIYQELTWKAAVGL					
PEKSKLQEIYQELTRKAAVGLPEKSKLQEIYQELTRKAAVGLPEKSKLQEIYQ					
WIKSIILAKKEVAGLVAVLRKAFENLQQLQSSSNKFTFMGSLSDNQTQWQVAGISPL					
SLPSKLVWNGEGLFNNVLELECAHESNQMNDQCNLAKFTQKRSASVSDPEQPL					
SAAPTPPPPPA"					

ORIGIN

Query Match 7.2%: Score 111; DB 6; Length 1312;
 Best Local Similarity 100.0%; Pred. No. 4,4e-52;
 Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 449 ATCTACGAGAGCTGACCTGATTAAGCTGATGCTGATGCTGATGAGGAATCCAG 508
 |||||||
 DB 270 ATCTACGAGAGCTGACCTGATTAAGCTGATGCTGATGCTGATGAGGAATCCAG 329
 |||||||

QY 509 CTGACGAGATCTACGAGAGCTGACCGAGCTGACAGCTGACAGTGGGCGAG 559
 |||||||
 DB 330 CTGACGAGATCTACGAGAGCTGACCGAGCTGACAGCTGACAGTGGGCGAG 580
 |||||||

RESULT 14
 HOMOLECTIN 1312 bp mRNA linear PR: 19-JUN-2000
 LOCUS Homo sapiens membrane-associated lectin type-C mRNA, complete cds
 DEFINITION M98457
 ACCESSION M98457.1 GI:187115
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eumetazoa; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE
 1 (bases 1 to 1312)
 Curtis, B.M., Scharnowski, S. and Watson, A.J.
 Sequence and expression of a membrane-associated C-type lectin that exhibits CD4-independent binding of human immunodeficiency virus envelope glycoprotein gp120
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 89 (17), 8356-8360 (1992)
 MEDLINE 92390446
 PUBMED 1518869

FEATURES

source
 1..1312
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="placenta"
 42..1256
 /codon_start=1
 /product="membrane-associated lectin type-C"
 /protein_id="AAE77072.1"
 /db_xref="GI:8572543"

CDS
 /translation="MSDSKEPRIGQILFEPTPLGIPROTGRKYSLAGGLGHPIV
 LLLSTLLAGLLVQVSKVPSISQSDRODAIYMLTOLKAAGVGLSKGLQETQ
 EITDIAAGVGLPEKSKIOETIYQELTRKAAGVGLPEKSKIOETIYQELTRKAAGVGL
 PEKSKIOETIYQELTRKAAGVGLPEKSKIOETIYQELTRKAAGVGLPEKSKIOETIYQ
 EITRKAAGVGLPEKSKIOETIYQELTRKAAGVGLPEKSKIOETIYQELTRKAAGVGL
 LTRKAAGVGLPEKSKIOETIYQELTRKAAGVGLPEKSKIOETIYQELTRKAAGVGL
 WHDSITACKKEVGAOLVIVKSAEDELFLQSSSRNFTWGLSDI NOESTWQVDSPL
 LPSFKQYMRGEPNNVNGEDCAEFSSNGMNDKCNLAKEWICKKSAASCSRDPEQL
 SPAATFNPPIPA"

BASE COUNT 336 a 348 c 372 g 256 t
 ORIGIN

Query Match 7.2%: Score 111; DB 6; Length 1312;
 Best Local Similarity 100.0%; Pred. No. 4,4e-52;
 Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 449 ATCTACGAGAGCTGACCTGATTAAGCTGATGCTGATGCTGATGAGGAATCCAG 508
 |||||||
 DB 270 ATCTACGAGAGCTGACCTGATTAAGCTGATGCTGATGCTGATGAGGAATCCAG 329
 |||||||

QY 509 CTGACGAGATCTACGAGAGCTGACCGAGCTGACAGCTGACAGTGGGCGAG 559
 |||||||
 DB 330 CTGACGAGATCTACGAGAGCTGACCGAGCTGACAGCTGACAGTGGGCGAG 580
 |||||||

RESULT 15
 AY042229 1315 bp mRNA linear PR: 28-AUG-2001
 LOCUS Homo sapiens mdc-sigmb type I isoform (CD209) mRNA, complete cds,
 alternatively spliced.

ACCESSION AY042229 GI:15281088
 VERSION AY042229.1 GI:15281088
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo sapiens.

REFERENCE
 1 (bases 1 to 1315)
 Mummidi, S., Catangay, J., Lam, L., Hoeltge, A., Tellos, J., Jimenez, F., Ahuja, S.S. and Ahuja, S.K.
 Extensive repertoire of membrane-bound and soluble cell-specific TCM-4-binding nonintegrin 1 (CD209) isoforms. Inter individual variation in expression transcripts

JOURNAL J. Biol. Chem. 276 (45), 42106-42112 (2001)
 MEDLINE 21413847
 PUBMED 11337487
 REFERENCE 2 (bases 1 to 1315)
 Mummidi, S. and Ahuja, S.K.
 Direct Submission
 Submitted (18-JUN-2001) Department of Medicine, Veterans Health Care System and University of Texas Center, 7703, Floyd Curl Drive, San Antonio, TX 78216

FEATURES

source
 1..1315
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="19"
 /map="19p11.4"
 1..1315
 /gene="CD209"
 1..1126
 /gene="CD209"
 /note="putative open reading frame; encoding protein tm2-sigmb"
 /evidence="not_experimental"
 101..1315
 /gene="CD209"
 /note="Similar to the nucleotide sequence Genbank Accession Number A008812; also /protein_id="AAK91854.1"
 /product="mdc-sigmb type I isoform"
 /codon_start=1
 /product="mdc-sigmb type I isoform"
 /protein_id="AAK91854.1"
 /db_xref="GI:15281089"

CDS
 /translation="MASACPSPTSHSEPTPLRQGL
 LLLSTLLAGLLVQVSKVPSISQSDRODAIYMLTOLKAAGVGLSKGLQETQ
 EITDIAAGVGLPEKSKIOETIYQELTRKAAGVGLPEKSKIOETIYQELTRKAAGVGL
 PEKSKIOETIYQELTRKAAGVGLPEKSKIOETIYQELTRKAAGVGLPEKSKIOETIYQ
 EITRKAAGVGLPEKSKIOETIYQELTRKAAGVGLPEKSKIOETIYQELTRKAAGVGL
 LTRKAAGVGLPEKSKIOETIYQELTRKAAGVGLPEKSKIOETIYQELTRKAAGVGL
 WHDSITACKKEVGAOLVIVKSAEDELFLQSSSRNFTWGLSDI NOESTWQVDSPL
 LPSFKQYMRGEPNNVNGEDCAEFSSNGMNDKCNLAKEWICKKSAASCSRDPEQL
 SPAATFNPPIPA"

BASE COUNT 332 a 345 c 496 g 252 t
 ORIGIN

Query Match 7.2%: Score 111; DB 9; Length
 Best Local Similarity 100.0%; Pred. No. 4,4e-52;
 Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 449 ATCTACGAGAGCTGACCTGATTAAGCTGATGCTGATGCTGATGAGGAATCCAG 508
 |||||||
 DB 327 ATCTACGAGAGCTGACCTGATTAAGCTGATGCTGATGCTGATGAGGAATCCAG 329
 |||||||

QY 509 CTGACGAGATCTACGAGAGCTGACCGAGCTGACAGCTGACAGTGGGCGAG 559
 |||||||
 DB 389 CTGACGAGATCTACGAGAGCTGACCGAGCTGACAGCTGACAGTGGGCGAG 580
 |||||||

RESULT 36
 AF290886 4266 bp mRNA linear PR: 28-AUG-2001
 LOCUS Homo sapiens mdc-sigmb type I isoform (CD209) mRNA, complete cds,
 alternatively spliced.

Db 1428 CTGACAGACATCTACACAGACCTGACCACTTAAGACTTCTACATGAGCTGAG 1478

RESULT 38
AC008812/c 143619 bp DNA linear PK1 28-JUL-2000

LOCUS AC008812
DEFINITION Homo sapiens chromosome 19 clone CTD-2102F19, complete sequence.
AC008812 GI:9558577
VERSION AC008812.7
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
COMMENT 2 (bases 1 to 143619)
DIRECT SUBMISSION
DOI: Joint Genome Institute.
JOURNAL Direct Submission
Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 143619)
DOE Joint Genome Institute and Stanford Human Genome Center.
REFERENCE Submitted (28-JUL-2000) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Jul 28, 2000 this sequence version replaced the
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 3.8.
FEATURES
Source
1. 143619
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTD-2102F19"

BASE COUNT 36628 a 45485 c 36414 g 36192 t

ORIGIN

Query Match 7.2%; Score 111; DB 9; Length 143619.
Best Local Similarity 100.0%; Pred No. 3 9e-52;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 449 ATCTACAGAACTGACCTGACCTTAAGCTGACAGTGGTGGTACCTGACAGAAATCCAAAG 508
|||||
Db 49148 ATCTACAGAACTGACCTGACCTTAAGCTGACAGTGGTGGTACCTGACAGAAATCCAAAG 49089
|||||

QY 509 CTGACAGACATCTACACAGACCTGACCACTTAAGACTTCTACATGAGCTGAG 559
|||||
Db 49088 CTGACAGACATCTACACAGACCTGACCACTTAAGACTTCTACATGAGCTGAG 49038
|||||

RESULT 39
AC008763/c 177962 bp DNA linear PK1 16-JUL-2000

LOCUS AC008763
DEFINITION Homo sapiens chromosome 19 clone CTD-3214H19, complete sequence.
AC008763
VERSION AC008763.9 GI:21844554
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
COMMENT 2 (bases 1 to 177962)

AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 177962)
DOE Joint Genome Institute and Stanford Human Genome Center.
REFERENCE Submitted (01-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
4 (bases 1 to 177962)
DOE Joint Genome Institute and Stanford Human Genome Center.
REFERENCE Submitted (16-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Jul 16, 2002 this sequence version replaced the
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.7% of Sequence;
Estimated Total Number of Errors is 0.5.
NOTE: Break in plasmid subclone coverage 122800-123177. Unresolved repeat copies 122800-123177.
FEATURES
Source
1. 177962
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTD-3214H19"
/note="NOTE: Break in plasmid subclone 122800-123177. Unresolved tandem repeat 122800-123177. Unresolved repeat copies 122800-123177." Location/Qualifiers

BASE COUNT 40697 a 49263 c 46202 g 41800 t

ORIGIN

Query Match 7.2%; Score 111; DB 9; Length 177962.
Best Local Similarity 100.0%; Pred No. 3 9e-52;
Matches 111; Conservative 0; Mismatches 0; Indels 0;

QY 449 ATCTACAGAACTGACCTGACCTTAAGCTGACAGTGGTGGTACCTGACAGAAATCCAAAG 508
|||||
Db 176849 ATCTACAGAACTGACCTGACCTTAAGCTGACAGTGGTGGTACCTGACAGAAATCCAAAG 508
|||||

QY 509 CTGACAGACATCTACACAGACCTGACCACTTAAGACTTCTACATGAGCTGAG 559
|||||
Db 176789 CTGACAGACATCTACACAGACCTGACCACTTAAGACTTCTACATGAGCTGAG 559
|||||

RESULT 40
AF369755 1146 bp mRNA linear PK1 2001

LOCUS AF369755
DEFINITION Macaca mulatta dendritic cell-specific TCAM-3 gene mRNA, complete cds.
AF369755
VERSION AF369755.1 GI:16118474
KEYWORDS
SOURCE Macaca mulatta.
ORGANISM Macaca mulatta.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca.
1 (bases 1 to 1146)
Hagarty, B.S., Ahmad, N., Morfitt, T., Edwards, J., Arnason, J., Reihardt, J.A., Kimata, J.T., Littman, D. and Doms, R.W.
FUNCTIONAL AND ANTIPODE CHARACTERIZATION OF HUMAN PLATINUM MACAQUE, AND MURINE PC-SIGN
J. Virol. 75 (21), 10281-10289 (2001)
JOURNAL 21465051
PUBMED 11581396

REFERENCE	Mosier W. and Stelzel J.M.
AUTHORS	
TITLE	Methods to identify polynucleotide and polypeptide sequences which may be associated with physiological and medical conditions
JOURNAL	Patent. WO 0146683-A 11 20-DEC-2001
FEATURES	Evolutionary Genomics, LLC (US)
SOURCE	location/Qualifiers
	1. 1212
	/organism="gorilla gorilla"
	/db_xref="taxon:9593"
BASE COUNT	315 a 319 c 349 g 229 t
ORIGIN	
Query Match	4.1%; Score 63; DB 6; Length 1212;
Best Local Similarity	100.0%; Pred. No. 2.7e-24;
Matches	63; Conservative 0; Mismatches 0; Indels 0 Gaps 0
QY	860 TGGCATTTCTTCACGCAAACTGTTACTTCATGATGTAAATGTCACAGCAATGCAATGAC 919
Db	778 TGGCAATTTCTTCACGCAAACTGTTACTTCATGATGTAAATGTCACAGCAATGCAATGAC 919
QY	920 TCC 922
Db	838 TCC 840
RESULT 44	
AF209480S1	603 bp DNA linear FR: 12 JAN-2001
LOCUS	AF209480S1
DEFINITION	Homo sapiens probable mannose binding C-type lectin DC-STK1 gene, exons 1 and 2.
ACCESSION	AF209480
VERSION	AF209480.1 GI:10179612
KEYWORDS	
SEGMENT	
SOURCE	1 of 2
ORGANISM	Homo sapiens.
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	Solileux, E.J., Barten, R. and Trowsdale, J.
AUTHORS	DC-STK1, a related gene, DC-STK1 and CD23 form a cluster on 19p13
TITLE	1 Immunol 145 (6) 2437-2442 (1996)
JOURNAL	20432267
MEDLINE	10975799
PUBMED	2 (bases 1 to 603)
REFERENCE	Solileux, E.J.
AUTHORS	Direct Submission
TITLE	Submitted (26 NOV-1999) Department of Pathology, University of
JOURNAL	Cambridge, Tennis Court Road, Cambridge, Cambs CB2 1QP, UK
FEATURES	location/Qualifiers
SOURCE	1. 603
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="19"
	/map_19p13; adjacent to CD23"
	1. 70
	/gene="abc"
	/number=1
EXON	151..234
EXON	/gene="abc"
	/number=2
BASE COUNT	144 a 123 c 218 g 118 t
ORIGIN	
Query Match	3.8%; Score 59; DB 9; Length 603;
Best Local Similarity	100.0%; Pred. No. 5.7e-22;
Matches	59; Conservative 0; Mismatches 0; Indels 0 Gaps 0
QY	257 CAGCGGCAAAACATCATGATCTTCAAGCAATCAAGCGTGAAGATGGGCTCTCTGGCG 315
Db	13 CAGCGGCAAAACATCATGATCTTCAAGCAATCAAGCGTGAAGCGTGGCGCTCTCTGGCG 315

```

RESULT 45          AF343727      1146 bp     mRNA       linear
LOCUS             AF343727      Macaca nemestrina dendritic cell-specific tcrm-3
DEFINITION        nonintegrin mRNA, complete cds.
ACCESSION         AF343727
VERSION           AF343727.1 GI:16116454
KEYWORDS
SOURCE            Macaca nemestrina
ORGANISM          Macaca nemestrina
REFERENCE         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Caprimulgidae; Macaca.
AUTHORS           Barilaud,F., Pohlman,S., Sparwasser,T., Yu,Kimata,J., Baur,G., Ahmad,N., Macfarlane,I., Edwards,T., Anason,J., Reinhardt,L.A., Kimata,J.T., Littner,I. and Doms,R.W.
TITLE             Functional and antigenic characterization of dendritic pituitary macaque, and murine DC-SIGN
JOURNAL           J. Virol. 75 (21), 10281-10289 (2001)
MEDLINE           11581396
PUBMED
REFERENCES        2 (basos, I to 1146)
AUTHORS           Kimata,J.T. and Yu,Kimata,M.T.
TITLE             Direct Submission
SUBMITTER         Submitted (29-JAN-2001) Virology & Immunology, School for Biomedical Research, Box 760549, San Antonio, TX 78245-0549, USA
FEATURES
   source          Location/Organisms
                     1..1146
                        /organism="Macaca nemestrina"
                        /db_xref="taxon:9545"
                        /cell_type="peripheral blood-derived dendritic cell"
                        /note="membrane-associated mannose binding protein"
CDS               /product="dendritic cell-specific tcrm-3 nonintegrin"
                   /protein_id="AAU14428.1"
                   /db_xref="GI:16116455"
                   /translation="MSDSEERFLQQLDEEGLGVGR,  
LTLSPFLLTAGLVVSKVPSTSGFSKODAYVMFTGK  
ETRLKAAGVELPEKSQOELDELTRLAAGELEPKSR,  
PEKRSQAIVGEIYELSRLKAAGVLPEKSQOELYORLTOKAT,  
LLOLKAAVERLCHECPMWITFPQGNCFMNSNRNHDSTLS,  
QNFLOVSRSRNGTMMWISLDNHGTGMVAOSPLPSFR,  
EFSSGNWDHDCNKLAFTCKNSASNSIDEEKLNPAP"  
BASE COUNT      296 a      313 c      328 g      214 t
ORIGIN
Query Match      3 7% Score 57; Dh 9; Length 1 ;  
Best Local Similarity 100.0%; Prod. No. 8,2e+21;  
Matches 57; Conservative 0; Mismatches 0; Indels 0
```

572 TCCACGCTCATGAGATTCAAGAGATGAAGCGCTGCATGG
|||||
Db 421 TCACAATTTCATGAGATTCAAGAGATGAAGCGCTGCATGG

4
.
.
.
.

RESULT 2	
AAV84361	AAV84361 standard; cDNA to mRNA. 1643 bp.
XX	
AC	AAV84361:
XX	
DT	30-MAP-1999 (first entry)
XX	
DE	Human liver cDNA clone HP01347.
XX	
KW	Transmembrane protein; HP01347; human; lectin; receptor; liver; ds
XX	
OS	Homo sapiens.
XX	
Key	location/Qualifiers
FH	25..915
FI	/*tag= a
FI	/note= "cDNA comprising the coding region (minus
FT	the stop codon) is claimed (claim 3)"
FT	
XX	

PN W09855508-A2.
XX
PD 10-DEC-1998.
XX
PE 03-JUN-1998; 98WO-1P02445.
XX
PH 03-JUN-1997; 97JP-014648.
XX
PA (PROJ-) PROTEGENE INC.
XX
PA (SAGA) SAGAMI CHEM RES CENTRE.
XX
PI Kato S, Sekine S, Yamaguchi T;
XX
XX WPI; 1999-045730/04.
DR P-PSDB; AAW88493.
XX
XX New human proteins containing transmembrane domains and their
PT encoding sequences - useful in the preparation of antibodies and
PT large-scale protein production, gene diagnosis, and gene therapy
XX
PS Claim 4; Page 125-126; 178pp; English.

CC This is the nucleotide sequence of cDNA clone H9D1347, which
CC includes a coding region (also claimed) for a novel human
CC transmembrane protein (see AA0884933). The clone was isolated from a
CC liver cDNA library using a signal sequence detection method and by
CC protein synthesis *in vitro* translation. The encoded protein
CC is characterised as a type-II membrane protein having a single
CC N-terminal transmembrane domain. It has 85.5% homology over 284
CC amino acid residues to the human HIV envelope glycoprotein gp120-
CC binding C-type lectin (a CD4-independent HIV receptor) and ap120-
CC function as a receptor on the membrane surface. The invention
CC provides nucleotide sequences (see AA043357-76) coding for 13
CC transmembrane proteins (see AA088491-508), vectors containing such
CC polynucleotides, and eukaryotic cells containing the vectors. The
CC proteins can be used as antigens or as compositions in the
CC preparation of antibodies against the proteins. The polynucleotides
CC can be used as probes for gene diagnosis, and as gene sources for
CC gene therapy and large-scale production of proteins encoded by the
CC cDNA. The host cells are used for the detection of ligands
CC corresponding to the expressed proteins, and the screening of low
CC mol.wt. medicines.

Query Match	33.6%	Score 518;	DB 20;	Length 1643;
Best Local Similarity	100.0%;	Pred. No. 6,5e-246;		
Matches 518;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

DY 257 CAGCGGGAACATCAGTGCATCCAAAGCAACCAGGTGGCACCTGGGCTTCCTGGGG 316
|||||
|||

Ddb 13 CAGCTGGAAAACATCATGTGACTCCAATTACAACAAAGTGTTTAGTAATCTAATACTCTCTGGGG 72
|||||
|||

OY	317	TGTTGGCATGGCCCTGAGTGGAACTGCTGCTTAACTGTCG
Db	73	TGTTGGCAATGAGTGGCCCTGAGTGGAACTGCTGCTTAACTGTCG
OY	377	CTGGTGGCATGCCCTGAGTGGAACTGCTGCTTAACTGTCG
Db	133	CTGGTGGCAATGCCCTGAGTGGAACTGCTGCTTAACTGTCG
OY	437	GAGCAAGCAATCTTACCAAAAGCTGCAAGTTCATACCTGCAATGGC
Db	193	GAGCAAGCAATCTTACCAAAAGCTGCAAGTTCATACCTGCAATGGC
OY	497	GATAAATCTAAATGATAGAAATCTAACTAAATGATGAAAGGCT
Db	253	GATAAATCTAAAGTGTAGAAATCTAACTAAATGATGAAAGGCT
OY	557	GAGTGGCAAGAAATCCAAATGAGTGGAACTGCTGCTTAACTGTCG
Db	313	GATGATGCAAGAAATCTAAAGTGTAGAAATCTGCTGCTTAACTGTCG
OY	617	GCATGAGTGGCACTTGCAGAAATAATCTAAAGTGTAGAAATCTATCTAGCA
Db	373	GCATGAGTGGCACTTGCAGAAATAATCTAAAGTGTAGAAATCTATCTAGCA
OY	677	CTGAAGAGCTGATGATGAGTGAATTGCTCAATAAAATCTAAAGTGTAGAAAT
Db	433	CTGAAGAGCTGATGATGAGTGAATTGCTCAATAAAATCTAAAGTGTAGAAAT
OY	737	CTGACCGGGGTGAAGGCTGCATGAGTGGAGTGGCCAA 774
Db	493	CTGACCGGGGTGAAGGCTGCATGAGTGGAGTGGCCAA 530

RESULT 3
AA165469 standard; DNA: 1643 bp.
XX
AC
XX
DT
10-DEC-2001 (first entry)
XX
DE
DNA encoding a splice variant of human DG-SIGN.

KW		Human receptor; DC-SIGN; dendritic cell; T lymphocyte
KM	qg120;	C-type lectin; ICAM3; HIV entry; "cell, macrophage"
KK		HIV infection; splice variant; ss.
XX		
OS	Homo sapiens.	
FH	Key	Location/Qualifiers
FT	CDS	25..915
PT		/tag= a
FT		/product "DC-SIGN splice variant"
XX		
PN	W0200164752-A2.	
XX		
PD	07-SEP-2001.	
XX		
FP	28-FEB-2001; 2001WC-US0642.	
XX		
EK	02-MAR-2000; 2000US-0517605.	
XX		
PA	(UYNY) UNIV NEW YORK STATE.	
PA	(UYNI-) UNIV NIJMEGEN.	
XX		
PI	Liltman DP, Kwon D, Van Kooyk Y, deLijndenbeek T;	
XX	WP1: 2001-602565/68.	
DR	P-PDB; AAG79093.	

PT An antibody for the treatment or prevention of HIV-infection binds to gp120 portion which binds to CD4 and CXCR4 receptors on target cells. The antibody is exposed in the presence of DC-SIGN due to concomitant conformational change.

RESULT 8

ABN94615/c

ID ABN94615 standard; DNA: 506 BP.

AC ABN94615;

DT 13-AUG-2002 (first entry)

DE Gene #1113 used to diagnose liver cancer.

KM Gene: liver cancer; ds: hepatocellular carcinoma; hepatotropic;
 KM metastatic liver tumor; cytostatic; expression profile; disease state;
 KM disease progression; drug toxicity; drug efficacy; drug metabolism.

OS Homo sapiens.

PN W020029103 A2.

PD 11-APR-2002.

PE 02-OCT-2001; 2001WO-0830589.

PR 02-OCT-2000; 2000US-237054P.

PA (GENE-) GENE LOGIC INC

PI Horne D, Alvares C, Peres Da-Silva S, Winkley TG.

PR 2002-426119/45.

PT Diagnosing and detecting the progression of liver cancer.

PT hepatocellular carcinoma or metastatic liver tumor in a patient,

PT involves detecting the level of expression of two or more genes in a

PT liver tissue sample.

PS Claim 1: SEQ ID NO 1113; 298pp; English.

CC The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumor in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN93503-ABN97455 in a
 CC tissue sample. The method of the invention has hepatotropic, and
 CC cytostatic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at http://wipo.int/pub/published_pat_sequences.

SQ Sequence 506 BP; 165 A; 112 C; 127 G; 102 T; 0 other;

Query Match 14.3%; Score 220; DB 24; Length 506;

Best Local Similarity 100.0%; Pred No 1 9e-98;

Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1318 CTTTGACTGGGGAGCTCTGTTGAAGAGCTTATCTTGACAGCTGGAGACACAGGGA 1377
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 1378 ATTAATTCGCCAGTCAATGATATCTAGAGAGATGAGAGATGAGAGATGAGAGAG 1437
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 1400 ATTAATTCGCCAGTCAATGATATCTAGAGAGATGAGAGATGAGAGATGAGAGAG 1501
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 1418 TTCCAGCCCGACATCTTCTTTGCTATACATGCTTCATTTGGCTTTTTCAGTTG 1497
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 1400 TTCCAGCCCGACATCTTCTTTGCTATACATGCTTCATTTGGCTTTTTCAGTTG 141
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 1498 TAGCTTATATAAAGTGTAATGTTGTAATGCAAAA 1537
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 40 TAGCTTATATAAAGTGTAATGTTGTAATGCAAAA 1

RESULT 9

ABA60904

ID ABA60904 standard; DNA: 592 BP.

AC ABA60904;

DT 01-FEB-2002 (first entry)

DE Human foetal liver single exon nucleic acid probe #9209

DE Human foetal liver single exon nucleic acid probe #9209

OS Homo sapiens.

PN W0200157277 A2.

PD 09-AUG-2001.

PE 30-JAN-2001; 2001WO-0509669.

PR 04-FEB-2000; 2000US-0186332.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0644066.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0238459.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

PR WPI: 2001-483447/52.

PT Human genome-derived single exon nucleic acid probes for

PT analyzing gene expression in human foetal liver.

PS Claim 1: SEQ ID NO 9209; 639pp - sequence listing; Eng.

CC The invention relates to a single exon nucleic acid probe
 CC measuring human gene expression in a sample derived from
 CC liver. The single exon nucleic acid probes may be used
 CC measuring and displaying gene expression in samples deri
 CC foetal liver. The present sequence is a single exon nuc
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form pa
 CC printed specification, but was obtained in electronic
 CC from WIPO at http://wipo.int/pub/published_pat_sequences.

SQ Sequence 592 BP; 154 A; 149 C; 168 G; 121 T; 0 other;

Query Match 9.9%; Score 153; DB 22; Length

Best Local Similarity 100.0%; Pred No 2 7e-75;

Matches 153; Conservative 0; Mismatches 0; Indels 0;

QY 830 GAAGCCCTGTGCGGCACTGTGTAAAGACATGACATTTCTTCAAGAGAG 1377
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 149 GAAGCCCTGTGCGGCACTGTGTAAAGACATGACATTTCTTCAAGAGAG 1377
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 890 ATGCTTAATCTCCAGAGGAGATGAGAGATGAGAGATGAGAGATGAGAGAG 1437
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 209 ATGCTTAATCTCCAGAGGAGATGAGAGATGAGAGATGAGAGATGAGAGAG 1501
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 950 CAGCTGCTGTAATCAAAATGTTGTAATGCAAAA 982
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 269 CAGCTGCTGTAATCAAAATGTTGTAATGCAAAA 301
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 10

XX
 PN WO200186003-A2.
 XX
 PD 15-MV-2001.
 XX
 PE 10-JAN-2001: 2001WO-US00665.
 XX
 PE 04-FEB-2000: 2000US-180312P.
 PE 16-MAY-2000: 2000US-207456P.
 PE 16-JUN-2000: 2000US-0608408.
 PE 03-AUG-2000: 2000US-064256.
 PE 21-SEP-2000: 2000US-234687P.
 PE 27-SEP-2000: 2000US-236359P.
 PE 04-OCT-2000: 2000US-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 P1 Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 D2 WPI: 2002-114183/15.
 XX
 PF Spatially-addressable set of single exon nucleic acid probes; used to
 PF measure gene expression in human lung samples -
 XX
 P3 claim 1: SEQ ID No 9539; 634pp; English.
 XX
 XX
 The invention relates to a spatially addressable set of single exon
 nucleic acid probes for measuring gene expression in a sample derived
 from human lung comprising single exon nucleic acid probes having one of
 12614 nucleic acid sequences mentioned in the specification, or their
 complements or the 12187 open reading frames derived from the 12614
 probes. Also included are a microarray comprising the novel set of
 probes; the novel set of probes which hybridize at high stringency to a
 nucleic acid expressed in the human lung; measuring gene expression in a
 sample derived from human lung, comprising (a) contacting the array with
 a collection of detectably labeled nucleic acids derived from human lung
 mRNA, and (b) measuring the label detectably bound to each probe of
 the array; identifying exons in a eukaryotic genome, comprising
 (a) algorithmically predicting at least one exon from genomic sequences
 of the eukaryote, and (b) detecting specific hybridisation of detectably
 labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 having a fragment identical to the predicted exon, the probe is included
 in the above mentioned microarray; assigning exons to a single gene,
 comprising (a) identifying exons from genomic sequence by the method
 above and (b) measuring the expression of each of the exons in several
 tissues and/or cell types using hybridisation to a single exon
 microarrays having a probe with the exon, where a common pattern of
 expression of the exons in the tissues and/or cell types indicates that
 the exons should be assigned to a single gene; a peptide comprising one
 of 12011 sequences, mentioned in the specification, or encoded by the
 probes/open reading frames (ORF). The probes are used for gene
 expression analysis, and for identifying exons in a gene, particularly
 using human lung derived mRNA and for the study of lung diseases
 such as asthma, lung cancer, chronic obstructive pulmonary disease
 (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 fibrosis, neurofibromatosis, tuberous sclerosis, Marfan's disease,
 Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 haemorrhoidosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 pulmonary alveolar proteinosis, Kartagener syndrome, the cystic
 pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 and myaline membrane disease. The present sequence is a single exon
 probe of the invention.
 Note: The sequence data for this patent did not form part
 of the printed specification, but was obtained in electronic
 format directly from WPI at
 http://wipo.int/pub/published/pct_sequences.

Query Match 9.98; Score 153; DB 24; Length 592;
 Best Local Similarity 100.0%; Pred No. 27e-65;
 Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 840 GAATCTCTGTCGTCGACCTGCTGAGGATGATATCTTCCAGGAACTGTACTTC 889
 DB 149 GAATCTCTGTCGTCGACCTGCTGAGGATGATATCTTCCAGGAACTGTACTTC 208
 QY 890 AATCTGAACTGTCGACCTGCTGAGGATGATATCTTCCAGGAACTGTACTTC 949
 DB 209 AATCTGAACTGTCGACCTGCTGAGGATGATATCTTCCAGGAACTGTACTTC 268
 QY 950 CAATCTCTGTCGTCGACCTGCTGAGGATGATATCTTCCAGGAACTGTACTTC 982
 DB 299 CAATCTCTGTCGTCGACCTGCTGAGGATGATATCTTCCAGGAACTGTACTTC 301
 RESULT 15
 ID ABA73421 standard; DNA; 152 BP.
 AC ABA73421;
 XX
 DT 01 FEB 2002 (first entry)
 XX
 TE Human foetal liver single exon nucleic acid probe #21726.
 XX
 KW Human foetal liver gene expression; single exon nucleic acid probe; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157277-A2.
 XX
 PD 09-AUG-2001.
 XX
 PE 30-JAN-2001: 2001WO-US00669.
 XX
 PE 04-FEB-2000: 2000US-0180312.
 PE 16-MAY-2000: 2000US-0207456.
 PE 16-JUN-2000: 2000US-0608408.
 PE 03-AUG-2000: 2000US-064256.
 PE 21-SEP-2000: 2000US-0234687.
 PE 27-SEP-2000: 2000US-0236359.
 PE 04-OCT-2000: 2000US-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 P1 Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 D2 WPI: 2001-483447/52.
 XX
 PF Human genome-derived single exon nucleic acid probes useful for
 PF analyzing gene expression in human foetal liver -
 XX
 P3 claim 4: SEQ ID No 21726; 639pp + sequence listing; English.
 XX
 XX
 The invention relates to a single exon nucleic acid probe for
 measuring human gene expression in a sample derived from human foetal
 liver. The single exon nucleic acid probes may be used for predicting,
 measuring and displaying gene expression in samples derived from human
 foetal liver. The present sequence is a single exon nucleic acid
 probe of the invention.
 Note: The sequence data for this patent did not form part of the
 printed specification, but was obtained in electronic format directly
 from WPI at http://wipo.int/pub/published/pct_sequences.

Query Match 9.98; Score 152; DB 22; Length 152;
 Best Local Similarity 100.0%; Pred No. 8.7e-65;
 Matches 152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 831 AATCTGAACTGTCGACCTGCTGAGGATGATATCTTCCAGGAACTGTACTTC 890
 DB 1 AATCTGAACTGTCGACCTGCTGAGGATGATATCTTCCAGGAACTGTACTTC 60
 QY 891 AATCTGAACTGTCGACCTGCTGAGGATGATATCTTCCAGGAACTGTACTTC 950


```

lb 121 AGCTGCTGATATCAAAACTGCTGAGGAG 152

RESULT 18
PAK48027
ID AAK48027 standard; DNA: 152 BP.
XX
XX AAK48027;
AC
XX 06-NOV-2001 (first entry)
XX
XX Human bone marrow expressed single exon probe SEQ ID NO: 22584.
XX
XX Human bone marrow expressed exon, gene expression analysis; probe;
XX microarray, cancer, leukemia, lymphoma, myeloma; ss.
XX
XX Homo sapiens.
XX
XX WQ200157276-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WQ-0500664.
XX
XX 04-FEB-2000; 2000US-0180312.
XX
XX 26-MAY-2000; 2000US-0207456.
XX
XX 30-JUN-2000; 2000US-0608408.
XX
XX 03-AUG-2000; 2000US-064446.
XX
XX 21-SEP-2000; 2000US-0234687.
XX
XX 27-SEP-2000; 2000US-0236359.
XX
XX 04-OCT-2000; 2000US-0024263.
XX
XX
XX (MOLF-) MOLECULAR DYNAMICS INC.
XX
XX Peim SG, Hanzel EK, Chen W, Park DE.
XX
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -
XX
XX Example 4, SEQ ID NO. 22584; 658bp; Sequence listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukemia and myeloma. The present sequence is one of
XX the probes of the invention.
XX
XX Sequence 152 BP; 37 A, 48 C, 38 G, 29 T, 0 other.
XX
Query Match 9 943 Score 152; DR 22; Length 152;
Best Local Similarity 100.0%; Pred No. 8.7e-65;
Matches 152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

lb 831 AACGGTGTGTCGACATCTGTAAAGACATGAAATTCCTTCMAAGAAATTCATTTCA 890
lb 1 AACGGTGTGTCGACATCTGTAAAGACATGAAATTCCTTCMAAGAAATTCATTTCA 890
lb 1 AACGGTGTGTCGACATCTGTAAAGACATGAAATTCCTTCMAAGAAATTCATTTCA 890
lb 891 TGTCTACTCCGACGACATCTGTAAAGACATGAAATTCCTTCMAAGAAATTCATTTCA 950
lb 61 TGTCTACTCCGACGACATCTGTAAAGACATGAAATTCCTTCMAAGAAATTCATTTCA 120
lb 951 AACTGTGTGTAATCAAAATCTGATGAGAGAG 982
lb 121 AGCTGCTGATATCAAAACTGCTGAGGAG 152

RESULT 19
AA153856
ID AA153856 standard; DNA: 152 BP.
XX
XX

```

```

AC AA153856;
XX
XX 17-OCT-2001 (first entry)
XX
XX Probe #22542 used to measure gene expression in human placenta sample.
XX
XX DE: Probe microarray; human, placenta; antenatal diagnosis;
XX probe; genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX WQ200157272-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WQ-0500663.
XX
XX 04-FEB-2000; 2000US-0180312.
XX
XX 26-MAY-2000; 2000US-0207456.
XX
XX 30-JUN-2000; 2000US-0608408.
XX
XX 03-AUG-2000; 2000US-0632366.
XX
XX 21-SEP-2000; 2000US-0234687.
XX
XX 27-SEP-2000; 2000US-0236359.
XX
XX 04-OCT-2000; 2000US-0024263.
XX
XX
XX (MOLF-) MOLECULAR DYNAMICS INC.
XX
XX Peim SG, Hanzel EK, Chen W, Park DE.
XX
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human placenta -
XX
XX Example 5, SEQ ID NO. 22542; 654bp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENPs).
XX The present sequence is one such probe. The probes are useful for
XX producing a microarray for predicting, measuring and displaying gene
XX expression in samples derived from human placenta. The probes are useful
XX for antenatal diagnosis of human genetic disorders.
XX
XX Sequence 152 BP; 37 A, 48 C, 38 G, 29 T, 0 other.
XX
Query Match 9 943 Score 152; DR 22; Length 152;
Best Local Similarity 100.0%; Pred No. 8.7e-65;
Matches 152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

lb 831 AACGGTGTGTCGACATCTGTAAAGACATGAAATTCCTTCMAAGAAATTCATTTCA 890
lb 1 AACGGTGTGTCGACATCTGTAAAGACATGAAATTCCTTCMAAGAAATTCATTTCA 890
lb 891 TGTCTACTCCGACGACATCTGTAAAGACATGAAATTCCTTCMAAGAAATTCATTTCA 950
lb 61 TGTCTACTCCGACGACATCTGTAAAGACATGAAATTCCTTCMAAGAAATTCATTTCA 120
lb 951 AACTGTGTGTAATCAAAATCTGATGAGAGAG 982
lb 121 AGCTGCTGATATCAAAACTGCTGAGGAG 152

RESULT 20
ABS21963
ID ABS21963 standard; DNA: 152 BP.
XX
XX
XX ABS21963;
AC
XX 19-AUG-2002 (first entry)
XX
XX Human genome-derived single exon probe CBF from lung SEQ ID NO: 21954.
XX
XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
XX chronic obstructive pulmonary disease; interstitial lung disease;
XX

```

KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW hermann-wertheim-sclerost; Gaucher's disease; Niemann-pick disease;
 KW hemmings-pudlak syndrome; sarcoidosis; Niemann haemolysis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagren syndrome;
 KW primary ciliary dyskinesia; fibrocystic pulmonary dysplasia;
 KW hyaline membrane disease; open reading frame; ORF.
 XX
 OS Homo sapiens.
 XX
 PN W0200186003-A2
 PN
 PD 15-NOV-2001.
 PD
 XX
 PF 30-JAN-2001; 2001WO-US000665.
 PF
 XX
 PR 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-063236P.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236353P.
 PR 04-OCT-2000; 2000GB-0024263.
 PR
 XX
 PA (MOLF-) MOLECULAR DYNAMICS INC.
 PA
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 TI WP1: 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 PT
 XX
 PS Claim 4; SEQ ID NO 21954; 63pp; English.
 PS
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one
 CC of 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes: the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung, measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene. A method comprising one
 CC of 12011 sequences, mentioned in the specification, or employed by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemolysis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagren syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a single exon
 CC probe open reading frame of the invention.
 CC
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic

[illegible]

01 AAM2563. The proteins can have activities based on the tissues and
 02 cells they are expressed in, such as: antineoplastic, antihemetic,
 03 antirheumatic, immunosuppressive; antibacterial; endocrine, cardiac,
 04 central nervous system, vitreous, anti HIV, fungicide, antitumor,
 05 cardiovascular, antidiabetic, antileptotic; haemostatic, vasodilatory;
 06 anticancer; osteoporosis; dermatologic; antiallergic; anasthenic;
 07 antidiabetic; cytotoxic; neuroprotective; antidepressant; neurotrophic;
 08 antiparasitic; and immunostimulant. The proteins and polypeptides
 09 encoding them can be used in gene therapy, antisense therapy and vaccine
 10 production. The proteins and polypeptides are useful for screening for
 11 agonists of antagonists of a protein and for the treatment and diagnosis
 12 of disorders associated with the activity of a protein e.g. inflammation,
 13 rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 14 neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 15 infections, autoimmune, genetic diseases, haemopoietic disorders,
 16 anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
 17 osteoporosis, severe combined immunodeficiency, eczema, allergic
 18 rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 19 Alzheimer's disease, Parkinson's disease, neurodegenerative and
 20 neurological disorders.

21 X Sequence 433 BP; 110 A; 114 G; 128 G; 81 T; 0 other;

22 Query Match 7.28; Score 111; DB 22; Length 433;
 23 Best Local Similarity 100.0%; Pred. No. 1.7e-44;
 24 Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

25 449 ATCTACCAAGACCTGACCTTAACCTGAGTGGTGTGTCGACAGAAATCCAG 508
 26 DB 225 ATCTACCAAGACCTGACCTTAACCTGAGTGGTGTGTCGACAGAAATCCAG 284
 27 509 CTGACAGAGATCTACCAAGAGTGAACCTGAGTGGTGTGTCGAG 559
 28 DB 285 CTGACAGAGATCTACCAAGAGTGAACCTGAGTGGTGTGTCGAG 335

RESULT 22

10 ABA91223 standard, cDNA, 1212 BP.

AC ABA91223;

DE 04-APR-2002 (first entry)

DE Human DC-SIGN cDNA coding region.

KE DC-SIGN; human; gene; genomics; evolution; AIDS; HIV 1; ss.

OS Homo sapiens.

PN W0200146603-A2

PN 20-DEC-2001.

PE 06-JUN-2001; 2001MO-US18410.

PE 09-JUN-2000; 2000US-0591435.

PE (EVOL-) EVOLUTIONARY GENOMICS LLC.

PE Messier W, Sikela JM.

PE WPI; 2002-130744/17.

PI Determining human or non-human primate polynucleotide or polypeptide
 PT sequences associated with a physiological trait and have undergone
 PT evolutionary changes, for therapeutic use, involves using statistical
 PT methods.

PS Example 19; Fig 11; 14pp; English.

CC The present sequence is that of the coding region of the human
 CC DC-SIGN gene. DC-SIGN is expressed on dendritic cells and is known

CC to provide a mechanism for transport of HIV-1 virus to the lymph
 CC nodes. HIV-1 binds to the extracellular portion of DC-SIGN and
 CC infects the undifferentiated T cells in the lymph nodes via their
 CC CD4 proteins. This ultimately leads to compromise of the immune
 CC system and to full-blown AIDS. The invention comprises a
 CC comparative genomics approach to identify specific gene changes
 CC responsible for differences in functions and diseases distinguishing
 CC humans from non-humans, particularly primates including gorilla,
 CC orangutan and especially chimpanzee. The evolutionary significance
 CC of a nucleotide change is determined by the ratio of the
 CC non-synonymous substitution rate (Ka) to the synonymous rate (Ks)
 CC of the nucleotide sequence. Polynucleotide and polypeptide
 CC sequences corresponding to evolved traits may be relevant to human
 CC diseases or conditions such as unique or enhanced human brain
 CC functions, longer human life spans, susceptibility or resistance to
 CC disease, including AIDS and cancer, and aesthetic traits such as
 CC hair growth. Ka/Ks ratios for chimpanzee/human, human/gorilla and
 CC chimpanzee/orilla DC-SIGN cDNA sequences are 1.3, 0.87 and 1.3,
 CC respectively. It is theorized that chimpanzee resistance to
 CC progression to full-blown AIDS may be due in part to inability of
 CC HIV-1 to bind to chimpanzee DC-SIGN for transport to the lymph
 CC nodes. After determining the 3 dimensional structure of DC-SIGN,
 CC a rational drug design approach can be used to mimic the effects
 CC of chimpanzee DC-SIGN without interfering with the normal functions
 CC of human DC-SIGN.

29 X Sequence 1212 BP; 314 A; 321 G; 348 G; 229 T; 0 other;

30 Query Match 7.28; Score 111; DB 24; Length 1212;
 31 Best Local Similarity 100.0%; Pred. No. 1.6e-44;
 32 Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

33 449 ATCTACCAAGACCTGACCTTAACCTGAGTGGTGTGTCGACAGAAATCCAG 508
 34 DB 229 ATCTACCAAGACCTGACCTTAACCTGAGTGGTGTGTCGACAGAAATCCAG 288
 35 509 CTGACAGAGATCTACCAAGAGTGAACCTGAGTGGTGTGTCGAG 559
 36 DB 289 CTGACAGAGATCTACCAAGAGTGAACCTGAGTGGTGTGTCGAG 339

RESULT 23

10 ABA91224 standard; cDNA; 1212 BP.

AC ABA91224;

DE 04-APR-2002 (first entry)

DE Chimpanzee DC-SIGN cDNA coding region.

KE DC-SIGN; chimpanzee; gene; genomics; evolution; AIDS; HIV-1; ss.

OS Pan troglodytes.

PN W0200196603-A2.

PN 20-DEC-2001.

PE 06-JUN-2001; 2001MO-US18410.

PE 09-JUN-2000; 2000US-0591435.

PE (EVOL-) EVOLUTIONARY GENOMICS LLC.

PE Messier W, Sikela JM.

PE WPI; 2002-130744/17.

PI Determining human or non human primate polynucleotide or polypeptide
 PT sequences associated with a physiological trait and have undergone
 PT evolutionary changes, for therapeutic use, involves using statistical
 PT methods.

disorders .

PS Claim 1, SEQ ID NO 5054: 1399bp + Sequence listing: English.

XX The invention relates to human polypeptides (AA179941-AA193841) and

XX the encoded proteins (AA179941-AA193841) that exhibit activity relating to

CC cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The

CC polypeptides and polypeptides are useful in gene therapy, vaccines or

CC peptide therapy. The polypeptides have various cytokine-like activities,

CC e.g. stem cell growth factor activity, hematopoiesis regulating

CC activity, tissue growth factor activity, immunomodulatory activity and

CC activity/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukemia, nervous system disorders, arthritis and

CC inflammation.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from Wipo

CC at file:wpofint/pub/published_pat_sequences.

XX Sequence 338 bp: 1st A, 5th C, 102 G, 68 T, 0 other.

XX Query Match: 2 59 Score 39, DB 22, Length 338;

XX Best Local Similarity: 100.0%, Freq No: 7, 1e-09;

XX Matches: 39; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

UY 1355 TTGAGAGCTGGAGAGCAGCAGCAATTAATTTGCGGCACTG 1393

DB 256 TTGAGAGCTGGAGAGCAGCAGCAATTAATTTGCGGCACTG 218

RESULT 32

AAS36709/c

ID AAS36709 standard. ENA. 32190 BP.

AC AAS36709.

DT 17-DEC-2001 (first entry)

XX Human cardiovascular system antigen genomic DNA SEQ ID NO 2209.

DE

XX Cardiovascular system antigen: human, mouse, rabbit, goat; horse; cat;

KW chicken, sheep, immunosuppressive, antiarthritis; vasotrophic; dog;

KW antineumatic; antiproliferative; cytoskeletal; cardiact; neuroprotective;

KW cerebroprotective; neurotrophic; antibacterial; virocid; fungicide; cancer;

KW ophthalmological; vulnarity; gene therapy; autoimmune disease; neoplasm;

KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;

KW cerebrovascular disorder; nervous system disorder; bacterial infection;

KW fungal infection; viral infection; ocular disorder; endocrine disorder;

KW gastrointestinal disorder; renal disorder; respiratory disorder;

KW wound healing; skin aging; organ transplantation; tissue regeneration;

KW anti-infectivity.

XX

XX Homo sapiens

OS

XX

XX MO200155321-A2.

PN

XX 02-AUG-2001

PD

XX

XX 17-JAN-2001. 2001W1-0801340

PR

XX 31-JAN-2000; 2000US-0176065

PR

XX 04-FEB-2000; 2000US-0180628

PR

XX 24-FEB-2000; 2000US-0184664

PR

XX 02-MAR-2000; 2000US-0184750

PR

XX 16-MAR-2000; 2000US-0184874

PR

XX 17-MAR-2000; 2000US-0190076

PR

XX 18-APR-2000; 2000US-0198123

PR

XX 19-MAY-2000; 2000US-0205515

PR

XX 07-JUN-2000; 2000US-0209467

PR

XX 28-JUN-2000; 2000US-0214886

PR

XX 30-JUN-2000; 2000US-0215135

PR

XX 07-JUL-2000; 2000US-0215647

PR

XX 07-JUL-2000; 2000US-0215880

PR 11-JUL-2000; 2000US-0217487

PR 11-JUL-2000; 2000US-0217496

PR 14-JUL-2000; 2000US-0218290

PR 26-JUL-2000; 2000US-0220663

PR 26-JUL-2000; 2000US-0220964

PR 14-AUG-2000; 2000US-0224518

PR 14-AUG-2000; 2000US-0224519

PR 14-AUG-2000; 2000US-0225213

PR 14-AUG-2000; 2000US-0225414

PR 14-AUG-2000; 2000US-0225566

PR 14-AUG-2000; 2000US-0225567

PR 14-AUG-2000; 2000US-0225758

PR 14-AUG-2000; 2000US-0225759

PR 14-AUG-2000; 2000US-0225759

PR 18-AUG-2000; 2000US-0226681

PR 22-AUG-2000; 2000US-0226681

PR 22-AUG-2000; 2000US-0226688

PR 22-AUG-2000; 2000US-0227182

PR 23-AUG-2000; 2000US-0227009

PR 30-AUG-2000; 2000US-0228424

PR 01-SEP-2000; 2000US-0229287

PR 01-SEP-2000; 2000US-0229343

PR 01-SEP-2000; 2000US-0229344

PR 01-SEP-2000; 2000US-0229345

PR 01-SEP-2000; 2000US-0229509

PR 05-SEP-2000; 2000US-0229513

PR 05-SEP-2000; 2000US-0229513

PR 06-SEP-2000; 2000US-0230437

PR 06-SEP-2000; 2000US-0230438

PR 08-SEP-2000; 2000US-0231242

PR 08-SEP-2000; 2000US-0231243

PR 08-SEP-2000; 2000US-0231244

PR 08-SEP-2000; 2000US-0231413

PR 08-SEP-2000; 2000US-0231414

PR 08-SEP-2000; 2000US-0232080

PR 08-SEP-2000; 2000US-0232081

PR 08-SEP-2000; 2000US-0232081

PR 12-SEP-2000; 2000US-02321968

PR 14-SEP-2000; 2000US-0232397

PR 14-SEP-2000; 2000US-0232398

PR 14-SEP-2000; 2000US-0232399

PR 14-SEP-2000; 2000US-0232400

PR 14-SEP-2000; 2000US-0232401

PR 14-SEP-2000; 2000US-0233063

PR 14-SEP-2000; 2000US-0233064

PR 14-SEP-2000; 2000US-0233065

PR 21-SEP-2000; 2000US-0234223

PR 21-SEP-2000; 2000US-0234274

PR 25-SEP-2000; 2000US-0234997

PR 25-SEP-2000; 2000US-0234998

PR 25-SEP-2000; 2000US-0234998

PR 25-SEP-2000; 2000US-0235484

PR 26-SEP-2000; 2000US-0235834

PR 27-SEP-2000; 2000US-0235834

PR 27-SEP-2000; 2000US-0235836

PR 29-SEP-2000; 2000US-0236327

PR 29-SEP-2000; 2000US-0236367

PR 29-SEP-2000; 2000US-0236368

PR 29-SEP-2000; 2000US-0236369

PR 29-SEP-2000; 2000US-0236370

PR 29-SEP-2000; 2000US-0236370

PR 02-OCT-2000; 2000US-0237037

PR 02-OCT-2000; 2000US-0237037

PR 02-OCT-2000; 2000US-0237039

PR 02-OCT-2000; 2000US-0237039

PR 02-OCT-2000; 2000US-0237040

PR 02-OCT-2000; 2000US-0237040

PR 13-OCT-2000; 2000US-0237945

PR 13-OCT-2000; 2000US-0237945

PR 20-OCT-2000; 2000US-0241785

PR 20-OCT-2000; 2000US-0241786

PR 20-OCT-2000; 2000US-0241787

PR 20-OCT-2000; 2000US-0241808

PR 20-OCT-2000; 2000US-0241808

CC specific for an ovarian tumour protein comprising contacting T cells
 CC with (II) or (III) is useful in design and preparation of
 CC ribozyme molecules for inhibiting expression of the tumour polypeptides
 CC and proteins in tumour cells; and to isolate a full length gene from a
 CC suitable library e.g., a tumour cDNA library using well known
 CC techniques.

XX Sequence 472 BP, 124 A, 110 C, 93 G, 145 T, 0 other.

Query Match 1.78; Score 26; FR 24; Length 472;
 Best Local Similarity 100.0%; Pred. No. 0.019;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1474 TTTCAATTTGGCTGTTCTGATGTA 1499
 DB 219 TTTCAATTTGGCTGTTCTGATGTA 194

RESULT 14
 AAD03249 standard; cDNA: 59 BP.
 AAD03249:
 13-JUN-2001 (first entry)

DE Mouse type II integral membrane protein, CIRE-like fragment #22 cDNA.
 XX
 KW Mouse; type II integral membrane protein; CIRE, antitumor vaccine;
 KW immunosuppressive; autoimmune disorder; allergy; viral infection; APC;
 KW antigen presenting cell; bacterial infection; macrophage; dendritic cell;
 KW DC; ss.
 XX
 OS Mus musculus.
 XX
 PN W0200119869-A1.
 XX
 PD 22-MAR-2001.
 PF 13-SEP-2000; 2000WO-AU01096.
 XX
 PR 13-SEP-1999; 99AU-0002788.
 XX
 PA (COON-) COUNCIL QUEENSLAND INST MEDICAL RES.
 XX
 PI Caminschi I, Wright MD, Shortman KD;
 XX
 DK WPI: 2001-244788/25.

PT Novel type II integral membrane protein, CIRE, preferentially expressed
 PT in dendritic cells, macrophages and their precursors stimulatory to T
 PT cells, useful for identifying compounds having immunoregulatory
 PT activity -
 XX
 PS Disclosure; Page 28; 53pp; English.

XX The present cDNA sequence encodes mouse CIRE-like fragment derived
 CC from a protein expressed on dendritic cells which has similar function
 CC to CIRE protein.
 CC
 CC The invention relates to mouse type II integral membrane protein, CIRE
 CC and its corresponding cDNA molecule which is preferentially expressed in
 CC dendritic cells, macrophages and their precursors stimulatory to T cells.
 CC CIRE is useful for screening immunoregulatory compounds and as a marker
 CC for dendritic cell (DC) population. CIRE specific ligand such as a
 CC monoclonal antibodies are useful for isolating an antigen presenting cell
 CC (APC) from a biological sample. CIRE DNA is also used as tools to analyse
 CC the properties and functions of the CIRE gene/protein. For example, the
 CC DNA molecules are used to generate animal models, which lack functional
 CC CIRE genes and to isolate regulatory regions of the CIRE gene. Such
 CC regulatory regions are used to selectively express exogenous genes in DC
 CC or APC. The invention also relates to methods for modulating an immune
 CC response in a subject by binding and inhibiting the function of an APC
 CC such as a myeloid DC. The CIRE-specific ligands are useful for modulating

CC immune response by interfering with the function, migration
 CC or DC or APC. The ligands bind to and interfere with the
 CC myeloid dendritic cells such that antigen processing is
 CC lymphoid DC which leads to immune suppression and anergy.
 CC Allergies and autoimmune disorders. The ligands are also
 CC or enhance the interaction of viruses or bacteria with the
 CC resistance to infection. These ligands are also used to
 CC such as vaccine components, to DC or APC.

XX Sequence 55 B, 3 A, 20 C, 13 G, 17 T, 0 other;

Query Match 1.58; Score 23; FR 22; Length 54;
 Best Local Similarity 100.0%; Pred. No. 0.6;
 Matches 23; Conservative 0; Mismatches 0; Indels 0;

QY 867 TCTTCAAGCAACTGTTACTTC 879
 DB 32 TCTTCAAGCAACTGTTACTTC 54

RESULT 35
 AAD03228
 ID AAD03228 standard; cDNA: 153 BP.
 XX
 AC AAD03228:
 13-JUN-2001 (first entry)

DE Mouse CIRE like fragment #1 encoding cDNA clone, RP23-45
 XX
 LE Mouse CIRE like fragment #1 encoding cDNA clone, RP23-45
 XX
 KW Mouse; type II integral membrane protein; CIRE; antitumor
 KW immunosuppressive; autoimmune disorder; allergy; viral
 KW antigen presenting cell; bacterial infection; macrophage
 KW DC; ss.
 XX
 OS Mus musculus.
 XX
 PN W0200119869-A1.
 XX
 PD 22-MAR-2001.
 PF 13-SEP-2000; 2000WO-AU01096.
 XX
 PR 13-SEP-1999; 99AU-0002788.
 XX
 PA (COON-) COUNCIL QUEENSLAND INST MEDICAL RES.
 XX
 PI Caminschi I, Wright MD, Shortman KD;
 XX
 DK WPI: 2001-244788/25.
 XX
 PT Novel type II integral membrane protein, CIRE, preferentially
 PT in dendritic cells, macrophages and their precursors stimulatory to
 PT cells, useful for identifying compounds having immunoregulatory
 PT activity -
 XX
 PS Disclosure; Page 24; 53pp; English.

XX The present cDNA sequence encodes CIRE-like fragment #1
 CC RP23-458N1. This fragment is derived from a protein expressed
 CC in dendritic cells which has similar function to CIRE protein.
 CC The invention relates to mouse type II integral membrane
 CC and its corresponding cDNA molecule which is preferentially
 CC expressed in dendritic cells, macrophages and their precursors stimulatory to
 CC CIRE is useful for screening immunoregulatory compounds

for dendritic cell (DC) population. C1RE specific ligand such as monoclonal antibodies are useful for isolating an antigen presenting cell (APC) from a biological sample. C1RE DNA is also used as tools to analyse the properties and functions of the C1RE gene/protein. For example, the C1RE DNA molecules are used to generate animal models, which lack functional C1RE genes and to isolate regulatory regions of the C1RE gene. Such regulatory regions are used to selectively express exogenous genes in DC or APC. The invention also relates to methods for modulating an immune response in a subject by binding and inhibiting the function of an APC such as a myeloid DC. The C1RE-specific ligands are useful for modulating immune response by interfering with the function, migration or maturation of DC or APC. The ligands bind to and interfere with the function of myeloid dendritic cells such that antigen processing is undertaken by lymphoid DC which leads to immune suppression and anergy for treating allergies and autoimmune disorders. The ligands are also useful to block or enhance the interaction of viruses or bacteria with DC, for improving resistance to infection. These ligands are also used to target molecules such as vaccine components, to DC or APC.

Query Match 1.5% Score 23, DB 23, Length 153;
Best Local Similarity 100.0%, Pred. No. 0.59;
Matches 23, Conservative 0, Mismatches 0, Indels 0, Gaps 0

867 TCTTCAGCAAGCACTTACTTC 889
|||||
38 TCTTCAGCAAGCACTTACTTC 60

RESULT 36
AAS80491/c
ID AAS80491 standard; cDNA: 2879 BP.

AC AAS80491;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #16295.

KM Human, chromosome mapping: gene therapy; forensic;

EV food supplement, medical imaging, diagnostic, genetic disorder, ss

XX Homo sapiens.

PN W0200175067-A2.

XX 11-OCT-2001.

PE 30-MAR-2001; 2001WO-0508631.

XX 31-MAR-2000; 2000US-0540217

PR 23-AUG-2000; 2000US-0649167.

XX (HYSK-) HYSKO INC.

PI Dmanac RT, Liu C, Tana YT;

XX WPI; 2001-639362/73.

DE P-PSDB: ABG16394.

XX New isolated polynucleotide and encoded polypeptides, useful in
PI diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PI biodiversity

XX Claim 1: SEQ ID No 16295, 103pp, English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantifying a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at IP wipo.int/patpublic/seq-1/seq-1/sequences.

Seq-ID: 2879 BP, 25, A, 279 C, 842 G, 517 T, 0 other;

Query Match 1.1%, Score 22, DB 23, Length 2879;
Best Local Similarity 100.0%, Pred. No. 1.8;
Matches 22, Conservative 0, Mismatches 0, Indels 0, Gaps 0

867 TCTTCAGCAAGCACTTACTTC 1493
|||||
2211 TCTTCAGCAAGCACTTACTTC 2190

RESULT 37
ABK34644/c
ID ABK34644 standard; cDNA: 2661 BP.

AC ABK34644;

DT 08-MAY-2002 (first entry)

DE Human cDNA for novel secreted protein, SEQ ID 413.

XX Human (ss) gene; secreted protein; immune deficiency; viral infection;

KM bacterial infection, fungal infection, autoimmune disorder, burn;

KM rheumatoid arthritis, multiple sclerosis, autoimmune thyroiditis;

KM diabetes, allergy, asthma, neurodegenerative disease, Crohn's disease;

KM Alzheimer's disease, Parkinson's disease, liver fibrosis, tumour;

KM coagulation disorder, haemophilia, inflammatory disorder, ulcer;

KM tissue regeneration; wound healing, hematopoiesis, myeloid deficiency;

XX Lymphoid cell deficiency.

XX Homo sapiens.

PN W0200177290 A2.

XX 18-OCT-2001.

PE 29-MAR-2001; 2001WO-0510295.

PR 06-APR-2000; 2000US-194941P.

XX (GENY) GENETICS INST INC.

PI Wong CC, Clark BF, Fochtel K, Argosino MJ, Howes SB, Resnik RJ;

XX Galkota K, Graham JR;

XX WPI; 2002-179422/23.

XX Six hundred and twenty five polynucleotides derived from a variety of
CC human tissue sources which encode secreted proteins, useful for
CC treating immune deficiencies and disorders such as autoimmune disorders
CC -
XX Claim 1: Page 213, 339pp, English.
XX The invention relates to 625 polynucleotides which have been derived from
CC a variety of human tissue sources and which encode novel secreted

CC proteins, their complements and sequences that hybridise to them.
 CC Also included are a vector comprising the polynucleotide, a host cell
 CC transformed with the vector, the proteins encoded by the
 CC polynucleotides, antibodies that bind to the proteins and identification
 CC of modulators of the proteins or the expression of the polynucleotide.
 CC The polynucleotides can be used as probes for the identification
 CC and isolation of full length cDNA and genomic DNA. The polynucleotides
 CC and proteins can also be used as nutritional supplements. The protein
 CC is useful in the treatment of various immune deficiencies and disorders
 CC such as viral infections, bacterial infections, fungal infections,
 CC autoimmune disorders (e.g. rheumatoid arthritis, multiple sclerosis,
 CC autoimmune thyroiditis and diabetes) and allergic reactions and
 CC conditions (e.g. asthma). They are also useful for treating
 CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's
 CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),
 CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also
 CC useful for tissue regeneration, for wound healing and in the treatment
 CC of burns, incisions and ulcers. The proteins are also useful for
 CC regulating haematopoiesis, for treating myeloid or lymphoid cell
 CC deficiencies. The present sequence is one of the 625 cDNA sequences
 CC encoding a secreted protein.

XX Sequence 2361 BP; 637 A, 535 C; 491 G, 697 T; 1 other;

Query Match 1.48; Score 21; DB 24; Length 2361.
 Best Local Similarity 100.0%; Pred. No. 5.6;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1479 TTGGCTGTTCTGAGTTGTA 1499
 ||||||||||||||||||||
 DB 852 TTGGCTGTTCTGAGTTGTA 832

RESULT 38

AAC30751/-
 ID AAC30751 standard; cDNA: 157 BP.

AC AAC30751;

XX 06-OCT-2000 (first entry)

DE Human secreted protein 5' EST, SEQ ID NO: 34826.

KW Human: 5' EST; expressed sequence tag, secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping. ss.

XX Homo sapiens.

FN EP103401-A2.

XX 06-SEP-2000.

FE 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US 0122487.

PA (GSEST) GENSET.

PI Dunas Milne Edwards J. Buchert A. Giordano J.;

DE WPI: 2000-503381/45.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

PS Claim 1: SEQ ID 34826; 71bp - CD-ROM. English

XX The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or poly(A) RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)

CC of the mRNA because they are often obtained from oligo-c
 CC libraries. Such ESTs are not well suited for isolating c
 CC derived from the 5' ends of mRNAs and even in those cases
 CC cDNA sequences have been obtained, the full 5' UTR is re
 CC 5' ESTs are derived from mRNAs with intact 5' ends and ca
 CC used to obtain full length cDNAs and genomic DNAs. 5' EST
 CC in diagnostic, forensic, gene therapy and chromosome map
 CC They are used to obtain upstream regulatory sequences and
 CC expression and secretion vectors.

XX Sequence 157 BP; 21 A; 62 C; 43 G; 31 T; 0 other;

Query Match 1.38; Score 20; DB 21; Length
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 20; Conservative 0; Mismatches 0; Indels 0;

QY 603 CCGGGCTGACGCTGCACTG 622
 ||||||||||||||||||||
 DB 26 CCGGGCTGACGCTGCACTG 7

RESULT 39

AAK59336/C
 ID AAK59336 standard; cDNA: 315 BP.

AC AAK59336;

XX 06-NOV-2001 (first entry)

DE Human immune/haematopoietic antigen encoding cDNA SEQ 1

XX Human: immune; haematopoietic; immune/haematopoietic ant

KW cytostatic; gene therapy; vaccine; metastasis. ss.

OS Homo sapiens.

PN W0200157182-A2.

PD 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US01364.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198124.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209447.

PR 28-JUN-2000; 2000US-0214886.

PR 01-JUN-2000; 2000US-0215145.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217486.

PR 14-JUL-2000; 2000US-0218240.

PR 26-JUL-2000; 2000US-0220047.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0225759.

RESULT 40
AAK75081/C
ID AAK75081 standard; DNA: 319 RP
XX AAK75081,
XX
XX 07-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen gene-wide sequence. SEQ ID NO: 25833
XX
XX Human; immune, haematopoietic; immune/haematopoietic antigen; cancer;
XX
XX cytostatic, gene therapy, vaccine; metastasis; ds
XX
XX Homo sapiens.
XX
XX W0200157182 A2.
XX
XX 09-AUG-2001
XX
XX 17-JAN-2001; 2001W01NS01454
XX
XX 41-JAN-2000; 2000US-0179665
XX
XX 04-FEB-2000; 2000US-0180628
XX
XX 24-FEB-2000; 2000US-0184664
XX
XX 02-MAR-2000; 2000US-0186350
XX
XX 16-MAR-2000; 2000US-0189874
XX
XX 17-MAR-2000; 2000US-0190076
XX
XX 18-APR-2000; 2000US-0198123
XX
XX 19-MAY-2000; 2000US-0205515
XX
XX 07-JUN-2000; 2000US-0209457
XX
XX 28-JUN-2000; 2000US-0214886
XX
XX 30-JUN-2000; 2000US-0215135
XX
XX 07-JUL-2000; 2000US-0216647
XX
XX 07-JUL-2000; 2000US-0216880
XX
XX 11-JUL-2000; 2000US-0217487
XX
XX 11-JUL-2000; 2000US-0217496
XX
XX 14-JUL-2000; 2000US-0218290
XX
XX 26-JUL-2000; 2000US-0220963
XX
XX 26-JUL-2000; 2000US-0220964
XX
XX 14-AUG-2000; 2000US-0224518
XX
XX 14-AUG-2000; 2000US-0224519
XX
XX 14-AUG-2000; 2000US-0225213
XX
XX 14-AUG-2000; 2000US-0225214
XX
XX 14-AUG-2000; 2000US-0225365
XX
XX 14-AUG-2000; 2000US-0225267
XX
XX 14-AUG-2000; 2000US-0225268
XX
XX 14-AUG-2000; 2000US-0225270
XX
XX 14-AUG-2000; 2000US-0225447
XX
XX 14-AUG-2000; 2000US-0225757
XX
XX 14-AUG-2000; 2000US-0225758
XX
XX 14-AUG-2000; 2000US-0225759
XX
XX 14-AUG-2000; 2000US-0226279
XX
XX 22-AUG-2000; 2000US-0226683
XX
XX 22-AUG-2000; 2000US-0226868
XX
XX 22-AUG-2000; 2000US-0227182
XX
XX 23-AUG-2000; 2000US-0227009
XX
XX 30-AUG-2000; 2000US-0228924
XX
XX 01-SEP-2000; 2000US-0229287
XX
XX 01-SEP-2000; 2000US-0229343
XX
XX 01-SEP-2000; 2000US-0229344
XX
XX 01-SEP-2000; 2000US-0229345
XX
XX 05-SEP-2000; 2000US-0229509
XX
XX 05-SEP-2000; 2000US-0229513
XX
XX 06-SEP-2000; 2000US-0230437
XX
XX 06-SEP-2000; 2000US-0230439
XX
XX 08-SEP-2000; 2000US-0231242
XX
XX 08-SEP-2000; 2000US-0231243
XX
XX 08-SEP-2000; 2000US-0231244
XX
XX 08-SEP-2000; 2000US-0231413
XX
XX 08-SEP-2000; 2000US-0231414
XX
XX 08-SEP-2000; 2000US-0232880

PR 08-SEP-2000; 2000US-0232081
PR 12-SEP-2000; 2000US-0231968
PR 14-SEP-2000; 2000US-0233397
PR 14-SEP-2000; 2000US-0233398
PR 14-SEP-2000; 2000US-0233399
PR 14-SEP-2000; 2000US-0233399
PR 14-SEP-2000; 2000US-0233400
PR 14-SEP-2000; 2000US-0233401
PR 14-SEP-2000; 2000US-0233663
PR 14-SEP-2000; 2000US-0233664
PR 14-SEP-2000; 2000US-0233665
PR 21-SEP-2000; 2000US-0234223
PR 21-SEP-2000; 2000US-0234274
PR 25-SEP-2000; 2000US-0234997
PR 25-SEP-2000; 2000US-0234998
PR 26-SEP-2000; 2000US-0235484
PR 27-SEP-2000; 2000US-0235834
PR 27-SEP-2000; 2000US-0235835
PR 29-SEP-2000; 2000US-0236327
PR 29-SEP-2000; 2000US-0236367
PR 29-SEP-2000; 2000US-0236368
PR 29-SEP-2000; 2000US-0236369
PR 29-SEP-2000; 2000US-0236370
PR 02-OCT-2000; 2000US-0236892
PR 02-OCT-2000; 2000US-0237037
PR 02-OCT-2000; 2000US-0237038
PR 02-OCT-2000; 2000US-0237039
PR 02-OCT-2000; 2000US-0237040
PR 13-OCT-2000; 2000US-0239335
PR 13-OCT-2000; 2000US-0239337
PR 20-OCT-2000; 2000US-0240560
PR 20-OCT-2000; 2000US-0241221
PR 20-OCT-2000; 2000US-0241785
PR 20-OCT-2000; 2000US-0241786
PR 20-OCT-2000; 2000US-0241787
PR 20-OCT-2000; 2000US-0241808
PR 20-OCT-2000; 2000US-0241809
PR 20-OCT-2000; 2000US-0241826
PR 01-NOV-2000; 2000US-0244517
PR 08-NOV-2000; 2000US-0246474
PR 08-NOV-2000; 2000US-0246475
PR 08-NOV-2000; 2000US-0246476
PR 08-NOV-2000; 2000US-0246477
PR 08-NOV-2000; 2000US-0246478
PR 08-NOV-2000; 2000US-0246523
PR 08-NOV-2000; 2000US-0246524
PR 08-NOV-2000; 2000US-0246525
PR 08-NOV-2000; 2000US-0246526
PR 08-NOV-2000; 2000US-0246527
PR 08-NOV-2000; 2000US-0246528
PR 08-NOV-2000; 2000US-0246532
PR 08-NOV-2000; 2000US-0246609
PR 08-NOV-2000; 2000US-0246610
PR 08-NOV-2000; 2000US-0246611
PR 08-NOV-2000; 2000US-0246613
PR 17-NOV-2000; 2000US-0249207
PR 17-NOV-2000; 2000US-0249213
PR 17-NOV-2000; 2000US-0249214
PR 17-NOV-2000; 2000US-0249215
PR 17-NOV-2000; 2000US-0249216
PR 17-NOV-2000; 2000US-0249217
PR 17-NOV-2000; 2000US-0249218
PR 17-NOV-2000; 2000US-0249244
PR 17-NOV-2000; 2000US-0249245
PR 17-NOV-2000; 2000US-0249264
PR 17-NOV-2000; 2000US-0249265
PR 17-NOV-2000; 2000US-0249297
PR 17-NOV-2000; 2000US-0249299
PR 17-NOV-2000; 2000US-0249300

PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251386.
 PR 05-DEC-2000; 2000US-0256714.
 PR 06-DEC-2000; 2000US-0251474.
 PR 08-DEC-2000; 2000US-0251854.
 PR 08-DEC-2000; 2000US-0251864.
 PR 08-DEC-2000; 2000US-0251864.
 PR 08-DEC-2000; 2000US-0251984.
 PR 08-DEC-2000; 2000US-0251984.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0250474.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PR Kosen CA, Barash SC, Ruben SM;
 PR WPI; 2001-483426/52
 PR Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PR useful for preventing, diagnosing and/or treating cancers and
 PR metastasis.
 PR Disclosures: SEQ ID NO 29893; 3071pp + Sequence listing; English
 PR
 PR AAK54951 to AAK64702 encod[ing] the human immune/hematopoietic antigen (1)
 PR amino acid sequences given in AAK62176 to AAK61921. (1) have protective
 PR activity, and can be used in gene therapy and vaccine production. (1)
 PR proteins and polynucleotides may be used in the prevention, diagnosis and
 PR treatment of diseases associated with inappropriate (1) expression. For
 PR example, they may be used to treat disorders associated with decreased
 PR expression by rectifying mutations or deletions in a patient's genome
 PR that affect the activity of (1) by expressing inactive proteins or to
 PR supplement the patient's own production of (1). Additionally, (1)
 PR polynucleotides may be used to produce the secreted (1), by inserting
 PR the nucleic acids into a host cell and culturing the cell to express the
 PR protein. (1) proteins and polynucleotides may be used to prevent,
 PR diagnose and treat immune/hematopoietic-related diseases, especially
 PR cancers and cancer metastases of hematopoietic-derived cells. AAK64704
 PR to AAK87694 represent human immune/hematopoietic antigen genetic
 PR sequences from the present invention. AAK54942 to AAK54950 and AAK64704
 PR represent sequences used in the exemplification of the present invention.
 PR
 PR Sequence 319 BP; 96 A; 77 G; 56 C; 60 T; 0 other;
 PR
 PR Query Match 1 3%, Score 20, FR 22; Length 319,
 PR Best Local Similarity 100.0%, Pred. No. 1st
 PR Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 PR
 PR UY 1480 TTGAGTGTTCGATGTTGTA 1499
 PR DB 238 TTGAGTGTTCGATGTTGTA 219
 PR
 PR RESULT 41
 PR AAK75082/
 PR ID AAK75082 standard; DNA; 319 BP.
 PR XX
 PR AC AAK75082.
 PR XX
 PR DI 07-NOV-2001 (first entry)
 PR XX
 PR DE Human immune/hematopoietic antigen genetic sequence SEQ ID NO:29894.
 PR XX
 PR KW Human immune/hematopoietic immune/hematopoietic antigen carrier;
 PR Cytotoxic; gene therapy; vaccine; metastasis; ds.
 PR XX
 PR OS Homo sapiens.
 PR XX
 PR PN W0200157162-A2.
 PR XX
 PR PD 09-AUG-2001

XX
 PR 17 JAN 2001; 2001US-02501354.
 PR
 PR 31 JAN 2000; 2000US-0179065.
 PR 04 FEB 2000; 2000US-0180128.
 PR 24 FEB 2000; 2000US-0184664.
 PR 02 MAR 2000; 2000US-0186350.
 PR 15 MAR 2000; 2000US-0189874.
 PR 17 MAR 2000; 2000US-0190076.
 PR 18 MAR 2000; 2000US-0198123.
 PR 19 MAR 2000; 2000US-0205515.
 PR 28 JUN 2000; 2000US-0204467.
 PR 30 JUN 2000; 2000US-0215135.
 PR 07 JUL 2000; 2000US-0216647.
 PR 07 JUL 2000; 2000US-0216880.
 PR 11 JUL 2000; 2000US-0217487.
 PR 11 JUL 2000; 2000US-0217496.
 PR 14 JUL 2000; 2000US-0218290.
 PR 26 JUL 2000; 2000US-0220963.
 PR 29 JUL 2000; 2000US-0220964.
 PR 14 AUG 2000; 2000US-0224518.
 PR 14 AUG 2000; 2000US-0224519.
 PR 14 AUG 2000; 2000US-0225213.
 PR 14 AUG 2000; 2000US-0225214.
 PR 14 AUG 2000; 2000US-0225266.
 PR 14 AUG 2000; 2000US-0225267.
 PR 14 AUG 2000; 2000US-0225268.
 PR 14 AUG 2000; 2000US-0225270.
 PR 14 AUG 2000; 2000US-0225447.
 PR 14 AUG 2000; 2000US-0225757.
 PR 14 AUG 2000; 2000US-0225758.
 PR 14 AUG 2000; 2000US-0225759.
 PR 18 AUG 2000; 2000US-0225779.
 PR 22 AUG 2000; 2000US-0226881.
 PR 22 AUG 2000; 2000US-0227868.
 PR 24 AUG 2000; 2000US-0227893.
 PR 30 AUG 2000; 2000US-0228924.
 PR 01 SEP 2000; 2000US-0229287.
 PR 01 SEP 2000; 2000US-0229343.
 PR 01 SEP 2000; 2000US-0229344.
 PR 01 SEP 2000; 2000US-0229345.
 PR 05 SEP 2000; 2000US-0229509.
 PR 05 SEP 2000; 2000US-0229513.
 PR 05 SEP 2000; 2000US-0230437.
 PR 05 SEP 2000; 2000US-0230438.
 PR 08 SEP 2000; 2000US-0231243.
 PR 08 SEP 2000; 2000US-0231244.
 PR 08 SEP 2000; 2000US-0231413.
 PR 08 SEP 2000; 2000US-0231414.
 PR 08 SEP 2000; 2000US-0232080.
 PR 08 SEP 2000; 2000US-0232081.
 PR 12 SEP 2000; 2000US-0231968.
 PR 14 SEP 2000; 2000US-0232397.
 PR 14 SEP 2000; 2000US-0232398.
 PR 14 SEP 2000; 2000US-0232399.
 PR 14 SEP 2000; 2000US-0232400.
 PR 14 SEP 2000; 2000US-0232401.
 PR 14 SEP 2000; 2000US-0233063.
 PR 14 SEP 2000; 2000US-0233064.
 PR 14 SEP 2000; 2000US-0233065.
 PR 21 SEP 2000; 2000US-0234223.
 PR 21 SEP 2000; 2000US-0234274.
 PR 25 SEP 2000; 2000US-0234997.
 PR 25 SEP 2000; 2000US-0234998.
 PR 26 SEP 2000; 2000US-0235484.
 PR 27 SEP 2000; 2000US-0235834.
 PR 27 SEP 2000; 2000US-0235836.
 PR 29 SEP 2000; 2000US-0236327.
 PR 29 SEP 2000; 2000US-0236327.
 PR 29 SEP 2000; 2000US-0236368.

```

PR 29-SEP-2000: 2000US-0236369.
PR 29-SEP-2000: 2000US-0236370.
PR 02-OCT-2000: 2000US-0236802.
PR 02-OCT-2000: 2000US-0237037.
PR 02-OCT-2000: 2000US-0237038.
PR 02-OCT-2000: 2000US-0237039.
PR 02-OCT-2000: 2000US-0237040.
PR 13-OCT-2000: 2000US-0239935.
PR 13-OCT-2000: 2000US-0239937.
PR 20-OCT-2000: 2000US-0240960.
PR 20-OCT-2000: 2000US-0241221.
PR 20-OCT-2000: 2000US-0241785.
PR 20-OCT-2000: 2000US-0241786.
PR 20-OCT-2000: 2000US-0241787.
PR 20-OCT-2000: 2000US-0241808.
PR 20-OCT-2000: 2000US-0241809.
PR 20-OCT-2000: 2000US-0241826.
PR 01-NOV-2000: 2000US-0244617.
PR 08-NOV-2000: 2000US-0244674.
PR 08-NOV-2000: 2000US-0244675.
PR 08-NOV-2000: 2000US-0244676.
PR 08-NOV-2000: 2000US-0244677.
PR 08-NOV-2000: 2000US-0244678.
PR 08-NOV-2000: 2000US-0244623.
PR 08-NOV-2000: 2000US-0244624.
PR 08-NOV-2000: 2000US-0244625.
PR 08-NOV-2000: 2000US-0244626.
PR 08-NOV-2000: 2000US-0244627.
PR 08-NOV-2000: 2000US-0244628.
PR 08-NOV-2000: 2000US-0244632.
PR 08-NOV-2000: 2000US-0244609.
PR 08-NOV-2000: 2000US-0244610.
PR 08-NOV-2000: 2000US-0244611.
PR 08-NOV-2000: 2000US-0244613.
PR 17-NOV-2000: 2000US-0249207.
PR 17-NOV-2000: 2000US-0249208.
PR 17-NOV-2000: 2000US-0249209.
PR 17-NOV-2000: 2000US-0249210.
PR 17-NOV-2000: 2000US-0249211.
PR 17-NOV-2000: 2000US-0249212.
PR 17-NOV-2000: 2000US-0249213.
PR 17-NOV-2000: 2000US-0249214.
PR 17-NOV-2000: 2000US-0249215.
PR 17-NOV-2000: 2000US-0249216.
PR 17-NOV-2000: 2000US-0249217.
PR 17-NOV-2000: 2000US-0249218.
PR 17-NOV-2000: 2000US-0249244.
PR 17-NOV-2000: 2000US-0249245.
PR 17-NOV-2000: 2000US-0249264.
PR 17-NOV-2000: 2000US-0249265.
PR 17-NOV-2000: 2000US-0249267.
PR 17-NOV-2000: 2000US-0249299.
PR 17-NOV-2000: 2000US-0249300.
PR 01-DEC-2000: 2000US-0250160.
PR 01-DEC-2000: 2000US-0250161.
PR 05-DEC-2000: 2000US-0251040.
PR 05-DEC-2000: 2000US-0251988.
PR 05-DEC-2000: 2000US-0256719.
PR 06-DEC-2000: 2000US-0251479.
PR 06-DEC-2000: 2000US-0251856.
PR 08-DEC-2000: 2000US-0251868.
PR 08-DEC-2000: 2000US-0251869.
PR 08-DEC-2000: 2000US-0251894.
PR 08-DEC-2000: 2000US-0251947.
PR 11-DEC-2000: 2000US-0254090.
PR 05-JAN-2001: 2001US-0254678.
XX
XX
PA (HIMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM.
XX
XX WPI: 2001-483426/52.
XX

```

```

PI Nucleic acids encoding human immune/haematopoietic anti-
PI useful for preventing, diagnosing and/or treating cancer
PI metastasis -
XX
XX
PS Disclosure: SEQ ID NO 25604: 307bp - sequence listing
XX
XX AAK4591 to AAK4702 encode the human immune/haematopoie-
XX amino acid sequences given in AAK4591 to AAK4702. (1) i-
XX activity, and can be used in gene therapy and vaccine pr-
XX proteins and polynucleotides may be used in the preventi-
XX treatment of diseases associated with inappropriate (1) i-
XX example, they may be used to treat disorders associated a-
XX expression by rectifying mutations or deletions in a pat-
XX that affect the activity of (1) by expressing inactive pr-
XX supplement the patient's own production of (1). Addition-
XX polynucleotides may be used to produce the secreted (1) i-
XX the nucleic acids into a host cell and culturing the cel-
XX protein. (1) proteins and polynucleotides may be used to
XX diagnose and treat immune/haematopoietic-related disease
XX cancer and cancer metastases of haematopoietic-derived
XX to AAK6764 represent human immune/haematopoietic anti-g-
XX sequences from the present invention. AAK4594 to AAK4599
XX represent sequences used in the exemplification of the in-
XX
XX Sequence 319 bp: 96 A; 77 C; 56 G; 90 T; 0 other:
XX
XX Query Match 1387 score 20; DB 22; Length 6
XX Best local Similarity 100.0%; Prod. No. 18;
XX Matches 20; Conservative 0; Mismatches 0; Indels
XX
XX QY 1480 HGGCGTTCAGACTGCA 1479
XX Db 236 HGGCGTTCAGACTGCA 219
XX
XX RESULT 42
XX AAK75083/C
XX ID AAK75083 standard, DNA: 319 bp.
XX
XX AC AAK75083;
XX
XX U1 07-NOV-2001 (first entry)
XX
XX DE Human immune/haematopoietic antigen genomic sequence SPL
XX
XX KW Human, immune, haematopoietic; immune/haematopoietic anti-
XX cytosstatic, gene therapy, vaccine, metastasis; ds.
XX
XX US Homo sapiens.
XX
XX IN Wc200157102 A2.
XX
XX PD 09-AUG-2001.
XX
XX XX 17-JAN-2001, 2001WO-0801354.
XX
XX 41-JAN-2000: 2000US-0179065.
XX 04-FEB-2000: 2000US-0180628.
XX 24-FEB-2000: 2000US-0184664.
XX 02-MAR-2000: 2000US-0186350.
XX 16-MAR-2000: 2000US-0189874.
XX 17-MAR-2000: 2000US-0190076.
XX 18-APR-2000: 2000US-0198123.
XX 07-MAY-2000: 2000US-0205515.
XX 07-JUN-2000: 2000US-0209467.
XX 28-JUN-2000: 2000US-0214886.
XX 30-JUN-2000: 2000US-0215135.
XX 07-JUL-2000: 2000US-0216647.
XX 07-JUL-2000: 2000US-0216880.
XX 11-JUL-2000: 2000US-0217487.
XX 14-JUL-2000: 2000US-0217496.
XX 26-JUL-2000: 2000US-0218290.
XX 26-JUL-2000: 2000US-0220963.

```

PR 25-JUL-2000 2000US-0226964
 PR 14-AUG-2000 2000US-0224518
 PR 14-AUG-2000 2000US-0224519
 PR 14-AUG-2000 2000US-0225213
 PR 14-AUG-2000 2000US-0225214
 PR 14-AUG-2000 2000US-0225264
 PR 14-AUG-2000 2000US-0225267
 PR 14-AUG-2000 2000US-0225268
 PR 14-AUG-2000 2000US-0225270
 PR 14-AUG-2000 2000US-0225447
 PR 14-AUG-2000 2000US-0225757
 PR 14-AUG-2000 2000US-0225758
 PR 14-AUG-2000 2000US-0225759
 PR 18-AUG-2000 2000US-0226279
 PR 22-AUG-2000 2000US-0226681
 PR 22-AUG-2000 2000US-0226688
 PR 22-AUG-2000 2000US-0227182
 PR 23-AUG-2000 2000US-0227069
 PR 30-AUG-2000 2000US-0228524
 PR 01-SEP-2000 2000US-0229287
 PR 01-SEP-2000 2000US-0229343
 PR 01-SEP-2000 2000US-0229344
 PR 01-SEP-2000 2000US-0229347
 PR 05-SEP-2000 2000US-0229509
 PR 05-SEP-2000 2000US-0229513
 PR 06-SEP-2000 2000US-0230437
 PR 06-SEP-2000 2000US-0230438
 PR 08-SEP-2000 2000US-0231242
 PR 08-SEP-2000 2000US-0231243
 PR 08-SEP-2000 2000US-0231244
 PR 08-SEP-2000 2000US-0231245
 PR 08-SEP-2000 2000US-0231413
 PR 08-SEP-2000 2000US-0231414
 PR 08-SEP-2000 2000US-0232080
 PR 09-SEP-2000 2000US-0232081
 PR 12-SEP-2000 2000US-0231468
 PR 14-SEP-2000 2000US-0232367
 PR 14-SEP-2000 2000US-0232368
 PR 14-SEP-2000 2000US-0232369
 PR 14-SEP-2000 2000US-0232400
 PR 14-SEP-2000 2000US-0232401
 PR 14-SEP-2000 2000US-0232403
 PR 14-SEP-2000 2000US-0234064
 PR 14-SEP-2000 2000US-0233065
 PR 21-SEP-2000 2000US-0234273
 PR 21-SEP-2000 2000US-0234274
 PR 25-SEP-2000 2000US-0234347
 PR 25-SEP-2000 2000US-0234348
 PR 26-SEP-2000 2000US-0234998
 PR 27-SEP-2000 2000US-0235834
 PR 27-SEP-2000 2000US-0235835
 PR 29-SEP-2000 2000US-0236327
 PR 29-SEP-2000 2000US-0236367
 PR 29-SEP-2000 2000US-0236368
 PR 29-SEP-2000 2000US-0236369
 PR 29-SEP-2000 2000US-0236370
 PR 02-OCT-2000 2000US-0236882
 PR 02-OCT-2000 2000US-0237037
 PR 02-OCT-2000 2000US-0237038
 PR 02-OCT-2000 2000US-0237039
 PR 02-OCT-2000 2000US-0237040
 PR 13-OCT-2000 2000US-0239995
 PR 13-OCT-2000 2000US-0239997
 PR 29-OCT-2000 2000US-0240000
 PR 29-OCT-2000 2000US-0241221
 PR 20-OCT-2000 2000US-0241785
 PR 20-OCT-2000 2000US-0241786
 PR 20-OCT-2000 2000US-0241787
 PR 20-OCT-2000 2000US-0241808
 PR 20-OCT-2000 2000US-0241809
 PR 20-OCT-2000 2000US-0241826
 PR 01-NOV-2000 2000US-0244517
 PR 08-NOV-2000 2000US-0246474
 PR 08-NOV-2000 2000US-0246475

PR 08-NOV-2000 2000US-0244515
 PR 08-NOV-2000 2000US-0246477
 PR 08-NOV-2000 2000US-0246478
 PR 08-NOV-2000 2000US-0246523
 PR 08-NOV-2000 2000US-0246524
 PR 08-NOV-2000 2000US-0246525
 PR 08-NOV-2000 2000US-0246526
 PR 08-NOV-2000 2000US-0246527
 PR 08-NOV-2000 2000US-0246528
 PR 08-NOV-2000 2000US-0246529
 PR 08-NOV-2000 2000US-0246599
 PR 08-NOV-2000 2000US-0246610
 PR 08-NOV-2000 2000US-0246611
 PR 08-NOV-2000 2000US-0246613
 PR 17-NOV-2000 2000US-0249207
 PR 17-NOV-2000 2000US-0249208
 PR 17-NOV-2000 2000US-0249209
 PR 17-NOV-2000 2000US-0249210
 PR 17-NOV-2000 2000US-0249211
 PR 17-NOV-2000 2000US-0249212
 PR 17-NOV-2000 2000US-0249213
 PR 17-NOV-2000 2000US-0249214
 PR 17-NOV-2000 2000US-0249215
 PR 17-NOV-2000 2000US-0249216
 PR 17-NOV-2000 2000US-0249217
 PR 17-NOV-2000 2000US-0249299
 PR 17-NOV-2000 2000US-0249300
 PR 17-NOV-2000 2000US-0249301
 PR 01-DEC-2000 2000US-0250391
 PR 01-DEC-2000 2000US-0250392
 PR 01-DEC-2000 2000US-0251030
 PR 01-DEC-2000 2000US-0251423
 PR 05-DEC-2000 2000US-0251671
 PR 06-DEC-2000 2000US-0251479
 PR 08-DEC-2000 2000US-0251856
 PR 08-DEC-2000 2000US-0251858
 PR 08-DEC-2000 2000US-0251859
 PR 08-DEC-2000 2000US-0251869
 PR 08-DEC-2000 2000US-0251989
 PR 08-DEC-2000 2000US-0251990
 PR 11-DEC-2000 2000US-0254097
 PR 05-JAN-2001 2000US-0259478

(JIBMA) HUMAN GENOME SCI INC.

Foster CA, Barash SC, Ruben SM.

WPL, 2001-48436/52.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis.

Disclosures: SEQ ID NO 29835, 30716 - Sequence listing: English.

AA654751 to AA654762 encode the human immune/hematopoietic antigen (1) amino acid sequences given by AA654751 to AA654762. (1) have cytoprotective activity and can be used in gene therapy and vaccine production. (1) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (1) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (1) by expressing inactive proteins or to supplement the patient's own production of (1). Additionally, (1) polynucleotides may be used to produce the secreted (1), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (1) proteins and polynucleotides may be used to prevent, diagnose and treat immune/hematopoietic-related diseases, especially cancers and cancer metastases of hematopoietic-derived cells. AA654763

CC to AAK67694 represent human immature/hematopoietic antigen genomic
CC sequences from the present invention AAK5442 to AAK5490 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
XX
SO Sequence 319 BP: 96 A; 77 C; 56 G; 90 T; 0 other;

Query Match 1.3%; Score 20; DB 22; Length 319;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1480 TTGCTGTTTGACTGTA 1499
|||||
DB 248 TTGCTGTTTGACTGTA 219

RESULT 44
AAK94106/c
ID AAK94106 standard; DNA: 692 BP.
XX
XX AAK94106;
XX
XX 23-MAY-2001 (first entry)
XX
XX Primer specific for DNA encoding secretory/membrane protein SHY ID 540.
XX
XX Human; secretory protein; membrane protein; vaccine; gene therapy;
XX rheumatoid arthritis; diabetes; PCR primer; SS.
XX
XX Synthetic.
XX
XX EPI067182-A2.
XX
XX 10-JAN-2001.
XX
XX 07-JUL-2000; 2000EP-0114090.
XX
XX 08-JUL-1999; 99JP-0194179.
XX 11-JAN-2000; 2000JP-0118775.
XX 02-MAY-2000; 2000JP-0183766.
XX
XX (HELIX) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;
XX
XX WPI: 2001-093989/11.
XX
XX Nucleic acids encoding secretory proteins/membrane proteins, useful in
XX gene therapy or as candidate target molecules in drug development -
XX
XX Claim 4; SEQ ID 540; 609pp + CD ROM; English.

CC This invention relates to nucleic acid sequences AAK93744 - AAK93916
CC which encode human secretory or membrane proteins represented by
CC AAK88317 - AAK88419 included in the invention are primers
CC AAK93917 - AAK94295 and AAK6232 - AAK6235 which are used to isolate the
CC cDNA sequences of the invention. The invention also includes methods for
CC the production of antibodies directed against the proteins, and cDNA
CC sequences, which can be used in vaccines. The polynucleotide sequences
CC can be used in gene therapy. The polynucleotide sequences and the
CC proteins they encode may be used in the prevention, treatment and
CC diagnosis of diseases associated with inappropriate secretory
CC protein/membrane protein expression. The nucleic acids and complementary
CC sequences may also be used as DNA probes in diagnostic assays
CC (e.g. polymerase chain reactions (PCR)) to detect and quantitate the
CC presence of similar nucleic acid sequences in samples. They may also be
CC used to study the expression and function of secretory proteins/membrane
CC polypeptides and their role in metabolism. The polypeptides may be used
CC as antigens in the production of antibodies against them and in assays to
CC identify modulators (agonists and antagonists) of expression and
CC activity. The antibodies and antagonists may also be used as therapeutic
CC agents to down regulate expression and activity. The antibodies may also
CC be used as diagnostic agents for detecting the presence of the
CC polypeptides in samples (e.g. by enzyme linked immunosorbant assay

CC (ELISA). Examples of diseases which may be treated include
CC arthritis and diabetes.
XX
XX
SO Sequence 692 BP: 149 A; 192 C; 192 G; 156 T; 3 other;

Query Match 1.3%; Score 20; DB 22; Length 692;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 0; Indels 0;
QY 366 TGGTGGGCTCTCGTGGCC 385
|||||
DB 223 TGGTGGGCTCTCGTGGCC 204

RESULT 44
AA087427
ID AA087427 standard; CDNA: 1983 BP.
XX
XX AA087427;
XX
XX 16-NOV-1995 (first entry)
XX
XX Rat GRK cDNA clone B24.
XX
XX
XX Polymerase chain reaction; PCR; amplification; primer; isolated
XX G protein-coupled receptor kinase; GRK6; catalytic domain
XX Rhodopsin kinase; KR; beta adrenergic receptor kinase; L
XX conserved amino acid; subdomain II; subdomain VII; graft
XX signal transduction; immune system; inflammation; human;
XX
XX Rattus rattus.
XX
XX
XX Key Location/Qualifiers
XX CDS 13..1743
XX FT /*tag= a
XX TT /product= GRK6
XX
XX W0450/460 A.
XX
XX 23-MAR-1995.
XX
XX 16-SEP-1994; 94W0-0S10447.
XX
XX 17-SEP-1993; 93US-0124942.
XX 31-MAR-1994; 94US-0221817.
XX
XX (ICOS-) ICOS cGRP.
XX
XX Chantry D, Gray PW, Hoekstra MF;
XX
XX WPI: 1995-131358/17.
XX P-PSDB: AAR71375.
XX
XX New G protein-coupled receptor kinase, GRK6, and correspond-
XX ing protein used in the production of GRK6 and for the immunisation, the
XX diagnosis of diseases of the immune system and inflammation.
XX
XX Example 3; Page 48-50; 71pp; English.

CC This sequence represents the cDNA clone, B24, which encodes
CC rat G protein-coupled receptor kinase, GRK6. This is derived
CC from a rat thymus library in lambda ZAP, using the PCR to
CC amplified from rat T cell cDNA using the primers given
CC The rat and human GRK6 sequences exhibit 95% identity at
CC level. Peptide fragments of GRK6 may be used in immunisation
CC and quantification of GRK6. They may also be used for the
CC activities of GRK6, esp. those involved in signal trans-
CC Modulators of GRK6 may be useful in treatment of disease
CC physiological conditions of the immune system, such as
CC
XX
XX Sequence 1983 BP: 453 A; 528 C; 596 G; 403 T; 0 other;
XX
XX
XX Query Match 1.3%; Score 20; DB 16; Length

Best local similarity: 100.0%, Prod. No. 17,
Matches: 20, Conservative: 0, Mismatches: 0, Indels: 0, Gaps: 0,

QY 562 CTAGAGACTGACCTGCTCA 611
|||||
DB 438 CCAAGAGCTGACCTGCTCA 457

RESULT 45
AA93916/c
ID AAF93916 standard; cDNA; 2646 BP.

AA93916;

23-MAY-2001 (first entry)

Human cDNA encoding a membrane or secretory protein clone pSEC0265.

Human; secretory protein; membrane protein; vaccinia gene therapy;

Rheumatoid arthritis; diabetes; ss.

Homo sapiens.

EP1067182-A2.

07-MAY-2000; 2000EP-0114090.

08-JUL-1999; 99JP-0194179.

11-JAN-2000; 2000JP-0118775.

02-MAY-2000; 2000JP-0183766.

(HELI-) HELIX RES INST.

Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;

WPI: 2001-093989/11.

P-PSDB: AAB88489.

Nucleic acids encoding secretory proteins/membrane proteins, useful in

gene therapy or as candidate target molecules in drug development.

Claim 1: SEQ ID 349; 609pp + CD ROM; English.

This invention relates to nucleic acid sequences AAF93744 - AAF93916

which encode human secretory or membrane proteins represented by

AA888317 - AAB88419. Included in the invention are primers

AA93917 - AAF94295 and AAF62332. AAF62335 which are used to isolate the

cDNA sequences of the invention. The invention also includes methods for

the production of antibodies directed against the proteins, and cDNA

sequences, which can be used in vaccines. The polynucleotide sequences

can be used in gene therapy. The polynucleotide sequences and the

proteins they encode may be used in the prevention, treatment and

diagnosis of diseases associated with inappropriate secretory

protein/membrane protein expression. The nucleic acids and complementary

sequences may also be used as DNA probes in diagnostic assays

(e.g. polymerase chain reactions (PCR)) to detect and quantitate the

presence of similar nucleic acid sequences in samples. They may also be

used to study the expression and function of secretory proteins/membrane

polypeptides and their role in metabolism. The polypeptides may be used

as antigens in the production of antibodies against them and in assays to

identify modulators (agonists and antagonists) of expression and

activity. The antibodies and antagonists may also be used as therapeutic

agents to down regulate expression and activity. The antibodies may also

be used as diagnostic agents for detecting the presence of the

polypeptides in samples (e.g. by enzyme linked immunosorbent assay

(ELISA)). Examples of diseases which may be treated include rheumatoid

arthritis and diabetes.

Best local similarity: 100.0%, Prod. No. 17,
Matches: 20, Conservative: 0, Mismatches: 0,

QY 346 TCGTCTGAGCTGCTGCTGCTG 385
|||||
DB 223 TCGTCTGAGCTGCTGCTGCTG 204

Search completed: December 7, 2002, 09:38:21
Job time: 348 secs

Query Match: 1.39; Score: 20; DB: 22; Length: 2646;

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: December 7, 2002, 06:12:23 : Search time 6.1 seconds

(without alignments)
7752.386 Million cell updates/sec

Title: US-09-831-458A-25

Percent score: 1542

Sequence: 1 atgaccacacgctgctgcat

Scoring table: OLIGO_NUC

Gapop 60.0, Capexr 60.0

Searched: 441362 seqs, 15338381 residues

Word size: 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2003000000

Post-processing: Listing first 45 summaries

Database: Issued_Patents_NA:
1: /cgn2_6/prodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/1/ina/5A_COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/prodata/1/ina/6A_COMB.seq.*
5: /cgn2_6/prodata/1/ina/6C_COMB.seq.*
6: /cgn2_6/prodata/1/ina/6C_COMB.seq.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	518	33.6	1643	4	US-09-517-605-16
2	111	7.2	1212	4	US-09-591-435-9
3	111	7.2	1212	4	US-09-591-435-10
4	111	7.2	1212	4	US-09-517-605-11
5	63	4.1	1212	4	US-09-591-435-11
6	20	1.3	1983	1	US-08-231-817-21
7	20	1.3	1983	1	US-08-454-413-21
8	20	1.3	1983	5	US-09-10487-21
9	19	1.2	301	4	US-09-222-575-29
10	18	1.2	177	2	US-09-012-684-20
11	18	1.2	5020	4	US-08-938-291A-3
12	17	1.1	324	4	US-09-064-411A-33
13	17	1.1	402	4	US-08-935-223-25
14	17	1.1	987	4	US-09-064-411A-43
15	17	1.1	936	4	US-09-064-411A-38
16	17	1.1	1098	4	US-09-064-411A-36
17	17	1.1	1140	4	US-09-064-411A-47
18	17	1.1	1164	4	US-09-538-871-3
19	17	1.1	1287	4	US-09-064-411A-35
20	17	1.1	1903	4	US-08-581-148C-15
21	17	1.1	2560	4	US-09-395-640-1
22	17	1.1	3388	4	US-09-141-206-1
23	17	1.1	3715	4	US-09-600-776-9
24	17	1.1	6605	4	US-08-769-309A-4
25	17	1.1	6605	3	US-08-924-570-4
26	17	1.1	6803	3	US-08-665-259-19
27	17	1.1	6803	3	US-08-762-500-19

28	17	1.1	11444	2	US-08-222-617A-26
29	17	1.1	12364	2	US-08-222-617A-1
30	17	1.1	111282	4	US-09-754-250-3
31	17	1.1	169998	4	US-09-676-610B-24
32	16	1.0	243	3	US-09-081-420-30
33	16	1.0	243	4	US-09-574-141A-30
34	16	1.0	243	4	US-09-707-780-30
35	15	1.0	368	4	US-08-411-507-22
36	15	1.0	368	4	US-07-965-875A-22
37	16	1.0	594	4	US-08-441-507-17
38	16	1.0	594	4	US-07-969-875A-17
39	16	1.0	662	4	US-08-441-507-1
40	16	1.0	662	4	US-07-969-875A-1
41	16	1.0	680	4	US-09-288-143-11
42	16	1.0	714	4	US-09-198-603C-26
43	16	1.0	759	4	US-08-441-507-20
44	16	1.0	759	4	US-07-969-875A-20
45	16	1.0	775	4	US-08-441-507-3

ALIGNMENTS

RESULT 1

US-09-517-605-16
Sequence 16, Application US/09517605

Patent No. 6391567

GENERAL INFORMATION:

APPLICANT: Liltman, Dan R.

APPLICANT: Kwon, Douglas S.

APPLICANT: van Kooyk, Yvette

APPLICANT: Geiltenbeck, Theo

TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF PEPTIDE

FILE REFERENCE: 1049-1-017

CURRENT APPLICATION NUMBER: US/09517605

CURRENT FILING DATE: 2000-03-02

NUMBER OF SEQ ID NOS: 17

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 16

LENGTH: 1643

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (25) .. (915)

US-09-517-605-16

Query Match

Best Local Similarity 33.6%

Score 518

DB 4

Length 1643

Prod. No 16-250

Matches 518

Conservative 0

Mismatches 0

Indels 0

Query

DB

Query

DB

Query

DB

Query

DB

Query

DB

Query

DB

Query

DB

Query

DB

Query

DB

Query

1 TITLE OF INVENTION: SEQUENCES WHICH MAY BE ASSOCIATED WITH PHYSIOLOGICAL
2 TITLE OF INVENTION: AND MEDICAL CONDITIONS
3 FILE REFERENCE: GENO. 200.2
4 CURRENT APPLICATION NUMBER: US/09/591.435
5 CURRENT FILING DATE: 2000-06-09
6 PRIOR APPLICATION NUMBER: 09/591.435
7 PRIOR FILING DATE: 2000-06-09
8 PRIOR APPLICATION NUMBER: 09/240.915
9 PRIOR FILING DATE: 1999-01-29
10 PRIOR APPLICATION NUMBER: 60/074.263
11 PRIOR FILING DATE: 1998-01-30
12 PRIOR APPLICATION NUMBER: 60/098.987
13 PRIOR FILING DATE: 1998-09-02
14 NUMBER OF SEQ ID NOS: 13
15 SOFTWARE: Patent In Ver. 2.0
16 SEQ ID NO: 11
17 LENGTH: 1212
18 TYPE: DNA
19 ORGANISM: Gorilla gorilla
20 US-09-591-435-11

Query Match 4.18; Score 63; DB 4; Length 1212;
Best Local Similarity 100.0%; Pred. No. 56-22;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 860 UGACATCTTCCAGGAGACTGTATCTATCTATCTCCAGGAGACTGGACGAC 919
DB 778 TGACATCTTCCAGGAGACTGTATCTATCTATCTCCAGGAGACTGGACGAC 837
UY 929 TCC 922
DB 838 TCC 840

RESULT 6
US-08-221-817-21
Sequence 21, Application US/08221817
Patent No. 5532151
GENERAL INFORMATION:
APPLICANT: Chantry, David
APPLICANT: Gray, Patrick W.
APPLICANT: Hoeckstra, Merle F.
TITLE OF INVENTION: A No. 5532151el G Protein-Coupled Receptor
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESS: Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/221,817
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/123,932
FILING DATE: 17 SEP 1993
ATTORNEY/AGENT INFORMATION:
NAME: No. 5532151and, Grete F.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31981
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856

1 INFORMATION FOR SEQ ID NO: 21:
2 SEQUENCE CHARACTERISTICS:
3 LENGTH: 1683 base pairs
4 TYPE: nucleic acid
5 STRANDEDNESS: single
6 TOPOLOGY: linear
7 MOLECULE TYPE: cDNA
8 FEATURE:
9 NAME/KEY: CDS
10 LOCATION: 13..1740
11 US-08-221-817-21

Query Match 1.38; Score 20; DB 1; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0;

UY 552 CTAGAGCTGACCTGGCTGA 911
DB 438 CTAGAGCTGACCTGGCTGA 457

RESULT 7
US-08-454-439-21
Sequence 21, Application US/08454439
Patent No. 5591618
GENERAL INFORMATION:
APPLICANT: Chantry, David
APPLICANT: Gray, Patrick W.
APPLICANT: Hoeckstra, Merle F.
TITLE OF INVENTION: A No. 5591618el G Protein-Coupled R
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESS: Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,439
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/221,817
FILING DATE: 31-MAR-1994
APPLICATION NUMBER: 08/123,932
FILING DATE: 17 SEP 1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5591618and, Grete F.
REGISTRATION NUMBER: 45,402
REFERENCE/DOCKET NUMBER: 31981
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 1983 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 13..1740

US-08-454-449-21

Query Match 1.3%; Score 20; DB 1; Length 1983;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

U7 592 C7AGGAGCTGACCGGCTGA 611
|||||
DB 438 C7AGGAGCTGACCGGCTGA 457

RESULT 8

PCT-US94-10487-21

Sequence 21 Application PC/TUS9410487
GENERAL INFORMATION:

APPLICANT: ICOS Corporation
TITLE OF INVENTION: A Novel G Protein Coupled Receptor
TITLE OF INVENTION: Kinase GRK6
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
STREET: 6300 Sears Tower, 213 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/10487

FILING DATE:
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/221,817

FILING DATE: 31 MAR 1994
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/123,932

FILING DATE: 17 SEP 1993
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Noland, Grete E.

REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 2766/31981
TELEPHONE: (312) 474-6400
TELEFAX: (312) 474-0448

TELEPHONE: (312) 474-6400
TELEFAX: (312) 474-0448

TELEPHONE: (312) 474-6400
TELEFAX: (312) 474-0448

TELEPHONE: (312) 474-6400
TELEFAX: (312) 474-0448

TELEPHONE: (312) 474-6400
TELEFAX: (312) 474-0448

TELEPHONE: (312) 474-6400
TELEFAX: (312) 474-0448

TELEPHONE: (312) 474-6400
TELEFAX: (312) 474-0448

TELEPHONE: (312) 474-6400
TELEFAX: (312) 474-0448

TELEPHONE: (312) 474-6400
TELEFAX: (312) 474-0448

TELEPHONE: (312) 474-6400
TELEFAX: (312) 474-0448

TELEPHONE: (312) 474-6400
TELEFAX: (312) 474-0448

TELEPHONE: (312) 474-6400
TELEFAX: (312) 474-0448

TELEPHONE: (312) 474-6400
TELEFAX: (312) 474-0448

RESULT 9

US-09-222-575-2970
Sequence 20 Application US/09222575

Patent No. 638747
GENERAL INFORMATION:

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

APPLICANT: Yaghi, Jiang
APPLICANT: Yaghi, Jiang

Query Match 1.2%; Score 19; DB 4; Length 301;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

U7 100 GAGGAGCTGACCGGCTGA 1078
|||||
DB 265 GAGGAGCTGACCGGCTGA 247

RESULT 10

US-09-032-684-20
Sequence 20 Application US/09032684

Patent No. 5882874
GENERAL INFORMATION:

APPLICANT: Fisher, Paul B.
TITLE OF INVENTION: PROTEIN DIFFERENTIAL

TITLE OF INVENTION: DISPLAY
NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP

STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/032,684
FILING DATE:

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:

NAME: White, John P.
REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 5551/JPW/AMG
TELEPHONE: (212) 391-0400
TELEFAX: (212) 391-0400

TELEPHONE: (212) 391-0400
TELEFAX: (212) 391-0400

TELEPHONE: (212) 391-0400
TELEFAX: (212) 391-0400

TELEPHONE: (212) 391-0400
TELEFAX: (212) 391-0400

TELEPHONE: (212) 391-0400
TELEFAX: (212) 391-0400

TELEPHONE: (212) 391-0400
TELEFAX: (212) 391-0400

TELEPHONE: (212) 391-0400
TELEFAX: (212) 391-0400

TELEPHONE: (212) 391-0400
TELEFAX: (212) 391-0400

TELEPHONE: (212) 391-0400
TELEFAX: (212) 391-0400

TELEPHONE: (212) 391-0400
TELEFAX: (212) 391-0400

Query Match 1.2%; Score 18; DB 2; Length 177;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

U7 592 C7AGGAGCTGACCGGCTGA 611
|||||
DB 438 C7AGGAGCTGACCGGCTGA 457

Best Local Similarity 100.0%; Pred. No. 20;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1525 TGTACTGCAAAAAAAA 1542
|||||
DB 155 TGTACTGCAAAAAAAA 172

RESULT 11

US-08-938-291A-3

Sequence 3, Application US/08938291A

Patent No. 6117673

GENERAL INFORMATION:

APPLICANT: Lev, Sima

APPLICANT: Plozman, Gregory D.

APPLICANT: Schlessinger, Joseph

TITLE OF INVENTION: RBCG PROTEINS AND RELATED

TITLE OF INVENTION: PRODUCTS AND METHODS

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

STREET: Suite 4700

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 MB

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: FastSeq

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/938,291A

FILING DATE: September 26, 1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/027,337

FILING DATE: October 11, 1996

ATTORNEY/AGENT INFORMATION:

NAME: Walburg, Richard J.

REGISTRATION NUMBER: 32,327

REFERENCE/DOCKET NUMBER: 228/172

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 5020 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: Linear

US-08-938-291A-3

Query Match 1 2% Score 18, DB 3, Length 5020;

Best Local Similarity 100.0%; Pred. No. 21;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1429 TACAACTCTTCACGCC 1446
|||||
DB 2344 TACAACTCTTCACGCC 2361

RESULT 12

US-09-064-411A-33/C

Sequence 33, Application US/09064411A

Patent No. 6331664

GENERAL INFORMATION:

APPLICANT: Rubin-Wilson, Beth

APPLICANT: Guo, Lining

APPLICANT: Skokut, Tom

APPLICANT: Young, Scott

APPLICANT: Folkerts, Otto

APPLICANT: Armstrong, Katherine

APPLICANT: Cowen, Neil M.

TITLE OF INVENTION: Nucleotide Sequences of Major Oligo-

TITLE OF INVENTION: Thioesterase and Palmitoyl-ACP Thio-

TITLE OF INVENTION: Their Use in the Regulation of Fat-

NUMBER OF SEQUENCES: 49

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dow Agrosciences Patent Department

STREET: 9330 Zionsville Road

CITY: Indianapolis

STATE: Indiana

COUNTRY: USA

ZIP: 46268

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25" disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Pat-Edit Pro, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/064,411A

FILING DATE: 22-APR-1998

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/045,827

FILING DATE: 05-MAY-1997

ATTORNEY/AGENT INFORMATION:

NAME: Horucki, Andrea T.

REGISTRATION NUMBER: 34651

REFERENCE/DOCKET NUMBER: 50444

TELECOMMUNICATION INFORMATION:

TELEPHONE: 317-337-4846

TELEFAX: 317-337-4847

INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:

LENGTH: 324 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..324

US-09-064-411A-33

Query Match 1.1% Score 17, DB 4, Length 324;

Best Local Similarity 100.0%; Pred. No. 63;

Matches 17; Conservative 0; Mismatches 0; Indels 4; Gaps 0;

QY 1114 CAGTGGAGACCAACA 1140
|||||
DB 100 CAGTGGAGACCAACA 84

RESULT 13

US-08-905-223-79

Sequence 79, Application US/08905223

Patent No. 6228929

GENERAL INFORMATION:

APPLICANT: Edwards, Jean-Baptiste D.

APPLICANT: Duclert, Aymeric

APPLICANT: Lacroix, Bruno

TITLE OF INVENTION: 5' PSIS FOR SECRETED PROTEINS

NUMBER OF SEQUENCES: 503

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe, Matveus, Olson & Bear

STREET: 501 West Broadway

CITY: San Diego

STATE: California

COUNTRY: USA

ZIP: 92101-3505

COMPUTER READABLE FORM:

```

MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: WIN95
SOFTWARE: WORD
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/905,723
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Israelson, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 402 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: CDNA
ORIGINAL SOURCE:
ORGANISM: Homo Sapiens
TISSUE TYPE: Brain
FEATURE:
NAME/KEY: slq.peptide
LOCATION: 61..372
IDENTIFICATION METHOD: Von Heijne matrix
OTHER INFORMATION: score 6.6
OTHER INFORMATION: seq HPSHVMFHPMA/00
US-09-905-223-79

Query Match: 1.1% Score 17; DB 4; Length 402;
Best Local Similarity 100.0%; Pred No 64;
Matches 17, Conservative 0, Mismatches 0, Indels 0, Gaps 0,

UY 264 AAACATGAGTACTGC 280
DB 56 AAACATGAGTACTGC 72

RESULT 14
US-09-064-411A-43/c
Sequence 43, Application US/09/064411A
Patent No. 6331664
GENERAL INFORMATION:
APPLICANT: Rubin-Wilson, Beth
APPLICANT: Guo, Lining
APPLICANT: Skokut, Tom
APPLICANT: Young, Scott
APPLICANT: Folkerts, Otto
APPLICANT: Armstrong, Katherine
APPLICANT: Cowen, Neil M.
TITLE OF INVENTION: Nucleotide Sequences of Maize Oleoyl-ACP
TITLE OF INVENTION: Thioesterase and Palmitoyl-ACP Thioesterase Genes and
NUMBER OF INVENTION: Their Use In The Regulation of Fatty Acid Content of Oil
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dow Agrosciences Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/064,411A
FILING DATE: 22-APR-1998

```

```

CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/045,827
FILING DATE: 05-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Potucki, Andrea I.
REGISTRATION NUMBER: 33651
REFERENCE/DOCKET NUMBER: 50433
TELECOMMUNICATION INFORMATION:
TELEPHONE: 417-337-4846
TELEFAX: 317-337-4847
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 987 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 11..970
US-09-064-411A-44

Query Match: 1.1% Score 17; DB 4; Length 987;
Best Local Similarity 100.0%; Pred No. 65;
Matches 17, Conservative 0, Mismatches 0, Indels 0, Gaps 0,

UY 1114 CAGTGGAGAGCCGACCA 1130
DB 257 CAGTGGAGAGCCGACCA 241

RESULT 15
US-09-064-411A-36/c
Sequence 36, Application US/09/064411A
Patent No. 6331664
GENERAL INFORMATION:
APPLICANT: Rubin-Wilson, Beth
APPLICANT: Guo, Lining
APPLICANT: Skokut, Tom
APPLICANT: Young, Scott
APPLICANT: Folkerts, Otto
APPLICANT: Armstrong, Katherine
APPLICANT: Cowen, Neil M.
TITLE OF INVENTION: Nucleotide Sequences of Maize Oleoyl-ACP
TITLE OF INVENTION: Thioesterase and Palmitoyl-ACP Thioesterase Genes and
NUMBER OF INVENTION: Their Use In The Regulation of Fatty Acid Content of Oil
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dow Agrosciences Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/064,411A
FILING DATE: 22-APR-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/045,827
FILING DATE: 05-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Potucki, Andrea I.
REGISTRATION NUMBER: 33651
REFERENCE/DOCKET NUMBER: 50433
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-337-4846

```

TELEFAX: 317-337-4847
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 996 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..996
US-09-064-411A-38
Query Match 1.1%; Score 17; DB 4; Length 996;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1114 CAGTGGAGAACCCACACA 1130
DB 283 CAGTGGAGAACCCACACA 267
RESULT 16
US-09-064-411A-36/C
Sequence 36, Application US/09064411A
Patent No. 6311664
GENERAL INFORMATION:
APPLICANT: Rubin-Wilson, Beth
APPLICANT: Guo, Lining
APPLICANT: Skokut, Tom
APPLICANT: Young, Scott
APPLICANT: Folkerts, Otto
APPLICANT: Armstrong, Katherine
TITLE OF INVENTION: Nucleotide Sequences of Maize Oleoyl-ACP
TITLE OF INVENTION: Thioesterase and Palmitoyl-ACP Thioesterase Genes and
TITLE OF INVENTION: Their Use In The Regulation of Fatty Acid Content of Oil
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dow Agrosciences Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/064,411A
FILING DATE: 22-APR-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/045,827
FILING DATE: 05-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Borucki, Andrea T.
REGISTRATION NUMBER: 33651
REFERENCE/DOCKET NUMBER: 50433
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-337-4846
TELEFAX: 317-337-4847
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 1098 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS

LOCATION: 1..1098
FEATURE:
NAME/KEY: mal_peptide
LOCATION: 103..1098
US-09-064-411A-36
Query Match 1.1%; Score 17; DB 4; Length 1098;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0;
QY 1114 CAGTGGAGAACCCACACA 1130
DB 385 CAGTGGAGAACCCACACA 369
RESULT 17
US-09-064-411A-47
Sequence 47, Application US/09064411A
Patent No. 6311664
GENERAL INFORMATION:
APPLICANT: Rubin-Wilson, Beth
APPLICANT: Guo, Lining
APPLICANT: Skokut, Tom
APPLICANT: Young, Scott
APPLICANT: Folkerts, Otto
APPLICANT: Armstrong, Katherine
TITLE OF INVENTION: Nucleotide Sequences of Maize Oleoyl-ACP
TITLE OF INVENTION: Thioesterase and Palmitoyl-ACP Thioesterase Genes and
TITLE OF INVENTION: Their Use In The Regulation of Fatty Acid Content of Oil
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dow Agrosciences Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/064,411A
FILING DATE: 22-APR-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/045,827
FILING DATE: 05-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Borucki, Andrea T.
REGISTRATION NUMBER: 33651
REFERENCE/DOCKET NUMBER: 50433
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-337-4846
TELEFAX: 317-337-4847
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 1140 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: YES
US-09-064-411A-47
Query Match 1.1%; Score 17; DB 4; Length 1140;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0;
QY 1114 CAGTGGAGAACCCACACA 1130
DB 1114 CAGTGGAGAACCCACACA 1130

1b 745 CAGTGGAGAGCCGACCA 751

RESULT 18

US-09-538-871-3
Sequence 3, Application US/09548871

Patent No. 6358718

GENERAL INFORMATION:

APPLICANT: Olson, Mark

APPLICANT: Szeleni, Alina

TITLE OF INVENTION: METHODS FOR STABILIZATION AND

FILE OF INVENTION: RENATURATION OF PROTEINS USING NUCLEOLAR PROTEINS P2

FILE REFERENCE: 13121.000202

CURRENT APPLICATION NUMBER: US/09/538,871

CURRENT FILING DATE: 2000-03-30

EARLIER APPLICATION NUMBER: 60/126,910

EARLIER FILING DATE: 1999-03-30

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 3

LENGTH: 1164

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE: OTHER INFORMATION: Description of artificial sequence: No 6358718c

US-09-538-871-3

Query Match 1.1% Score 17; DB 4; Length 1164;

Best Local Similarity 100.0%; Prev. No. 65;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1454 CTTGAGAGTTGTAAGCA 1370

1b 992 CTTGAGAGTTGTAAGCA 1098

RESULT 19

US-09-064-411A-35/c
Sequence 35, Application US/09064411A

Patent No. 6331664

GENERAL INFORMATION:

APPLICANT: Rudin-Wilson, Beth

APPLICANT: Guo, Lining

APPLICANT: Skokul, Tom

APPLICANT: Young, Scott

APPLICANT: Folkeerts, Otto

APPLICANT: Armstrong, Katherine

TITLE OF INVENTION: Nucleotide Sequences of Maize Oleoyl-ACP

TITLE OF INVENTION: Thioesterase and Palmitoyl-ACP Thioesterase Genes and

NUMBER OF SEQUENCES: 49

CORRESPONDENCE ADDRESS:

ADDRESS: Dow AgroSciences Patent Department

STREET: 9330 Zionsville Road

CITY: Indianapolis

STATE: Indiana

COUNTRY: USA

ZIP: 46268

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/064,411A

FILING DATE: 22-APR-1998

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/045,827

FILING DATE: 05-MAY-1997

ATTORNEY/AGENT INFORMATION:

NAME: Hornicki, Andrea T.

REGISTRATION NUMBER: 33651

REFERENCE/DOCKET NUMBER: 50433

TELECOMMUNICATION INFORMATION:

TELEPHONE: 317-347-4846

TELEFAX: 317-347-4847

INFORMATION FOR SEQ ID NO: 45;

SEQUENCE CHARACTERISTICS:

LENGTH: 1287 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

FEATURE:

NAME/KEY: CUS

LOCATION: 38..1135

FEATURE:

NAME/KEY: mal_peptide

LOCATION: 140..1135

US-09-064-411A-35

Query Match 1.1% Score 17; DB 4; Length 1287;

Best Local Similarity 100.0%; Prev. No. 65;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1114 CAGTGGAGAGCCGACCA 1130

1b 422 CAGTGGAGAGCCGACCA 406

RESULT 20

US-08-581-148C-15/c

Sequence 15, Application US/08581148C

Patent No. 6060644

GENERAL INFORMATION:

APPLICANT: Schmale, Patrick S.

APPLICANT: Robertson, Donald S.

APPLICANT: Hansen, Joel D.

APPLICANT: Nikolau, Basil J.

APPLICANT: Xu, Xiaojie

TITLE OF INVENTION: ISOLATION AND USE OF CUTICULAR LIPID

TITLE OF INVENTION: GENES

NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:

ADDRESS: Levitt, Veit & Mayer, Ltd.

STREET: Two Prudential Plaza, Suite 4900

CITY: Chicago

STATE: IL

COUNTRY: US

ZIP: 60601-6780

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/581,148C

FILING DATE: 29-DEC-1995

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Lartzer, Carol

REGISTRATION NUMBER: 35243

REFERENCE/DOCKET NUMBER: 71380

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 616-5600

TELEFAX: (312) 616-5700

INFORMATION FOR SEQ ID NO: 15;

SEQUENCE CHARACTERISTICS:

LENGTH: 1903 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLBIOLE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1665
US-08-581-148C-15

Query Match 1.18; Score 17; DB 3; Length 1903;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
UY 43 GCAGCTCTTCCTCTGT 49
|||||
DB 567 GCAGCTCTTCCTCTGT 551

RESULT 21
US-09-305-640-1
Sequence 1, Application US/09305640B
Patent No. 6255468
GENERAL INFORMATION:
APPLICANT: SmithKline Beecham plc
TITLE OF INVENTION: No. 6255468-1 Compounds
FILE REFERENCE: GP30124
CURRENT APPLICATION NUMBER: US/09/305,640B
CURRENT FILING DATE: 1999-05-05
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 1
LENGTH: 2560
TYPE: DNA
ORGANISM: Homo sapiens
US-09-305-640-1

Query Match 1.18; Score 17; DB 4; Length 2560;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
UY 844 CCAGCTGCCAAGGACT 860
|||||
DB 2499 CCAGCTGCCAAGGACT 2515

RESULT 22
US-09-141-206-1
Sequence 1, Application US/09141206
Patent No. 6187559
GENERAL INFORMATION:
APPLICANT: Steed, Paul M
APPLICANT: Lasala, Daniel J.
TITLE OF INVENTION: Amino Acid Sequence of Human PLD2A
Patent No. 6187559
FILE REFERENCE: 4-30148/P1/091954/P
CURRENT FILING DATE: 1998-08-27
CURRENT FILING DATE: 1998-08-27
EARLIER FILING DATE: 1997-08-28
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 3388
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: CDS
LOCATION: (162)..(2963)
OTHER INFORMATION: Human PLD2
US-09-141-206-1

Query Match 1.18; Score 17; DB 4; Length 3388;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
UY 981 AACTCTACAGCTGCA 999

DB 1107 AACTCTACAGCTGCA 1124
|||||

RESULT 23
US-09-600-776-9/c
Sequence 9, Application US/09600776
Patent No. 6326168
GENERAL INFORMATION:
APPLICANT: Yamaguchi Pharmaceutical Co., Ltd.
TITLE OF INVENTION: A novel potassium channel protein
FILE REFERENCE: 19903-PCF
CURRENT APPLICATION NUMBER: US/09/600,776
CURRENT FILING DATE: 2000-07-21
PRIOR APPLICATION NUMBER: JP P1998-011444
PRIOR FILING DATE: 1998-01-23
PRIOR APPLICATION NUMBER: JP P1998-446198
PRIOR FILING DATE: 1998-12-04
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 3715
TYPE: DNA
ORGANISM: Rattus sp.
US-09-600-776-9

Query Match 1.18; Score 17; DB 4; Length 37
Best Local Similarity 100.0%; Pred. No. 68;
Matches 17; Conservative 0; Mismatches 0; Indels 0;

UY 919 CTCCTCACCCGCTGCG 935
|||||
DB 2825 CTCCTCACCCGCTGCG 2809

RESULT 24
US-08-769-309A-4/c
Sequence 4, Application US/08769309A
Patent No. 5741890
GENERAL INFORMATION:
APPLICANT: Scott, John J.
APPLICANT: Nauert, Brian J.
APPLICANT: Klaucek, Theresa M.
TITLE OF INVENTION: Protein Binding Domains of G-protein
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, C/Taylor, Gervasio, Murray & Box
STREET: 6300 Sears Tower/233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/769,309A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5741890and, Giga E.
REGISTRATION NUMBER: 45,402
REFERENCE/DOCKET NUMBER: 27866/34451
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6605 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 192..5531
US-08-769-409A-4

Query Match 1.1%; Score 17; DB 1; Length 6605;
Best Local Similarity 100.0%; Prod. No. 69;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 367 GGTGGAGTCTCTGCG 383
DB 846 GGTGGAGTCTCTGCG 830

RESULT 25
US-08-994-570-4/c
Sequence 4, Application US/08994570
Patent No. 6090929
GENERAL INFORMATION:

APPLICANT: Scott, John D.
APPLICANT: Nauer, Brian J.
APPLICANT: Klauk, Theresa M.
TITLE OF INVENTION: Protein Binding Domains of Gravin
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borum
STREET: 6300 Sears Tower/233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/994,570
FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: No. 6090929and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/4441
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 6605 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:

NAME/KEY: CDS
LOCATION: 192..5531
US-08-994-570-4

Query Match 1.1%; Score 17; DB 3; Length 6605;
Best Local Similarity 100.0%; Prod. No. 69;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 367 GGTGGAGTCTCTGCG 383
DB 846 GGTGGAGTCTCTGCG 830

RESULT 26

US-08-665-259-19/c
Sequence 19, Application US/08665259
Patent No. 6028173
GENERAL INFORMATION:

APPLICANT: Landes, Gregory M.
APPLICANT: Burn, Timothy C.
APPLICANT: Connors, Timothy D.
APPLICANT: Dackowski, William R.
APPLICANT: Van Raay, Terence J.
TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
METHODS OF MAKING AND USING SAME
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESS: GENZYME CORPORATION
STREET: One Mountain Road
CITY: Framingham
STATE: Massachusetts
COUNTRY: United States of America
ZIP: 01701

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/665,259
FILING DATE: 17-JUN-1996
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Dugan, Deborah A.
REGISTRATION NUMBER: 17,415
REFERENCE/DOCKET NUMBER: 165-9.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508) 872-8400
TELEFAX: (508) 872-5415

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:
LENGTH: 6803 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-665-259-19

Query Match 1.1%; Score 17; DB 3; Length 6803;

Best Local Similarity 100.0%; Prod. No. 69;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 398 GGTGCAAGGTCTCCAG 414
DB 4097 GGTGCAAGGTCTCCAG 4081

RESULT 27

US-08-762-500-19/c
Sequence 19, Application US/08762500
Patent No. 6030806
GENERAL INFORMATION:

APPLICANT: Landes, Gregory M.
APPLICANT: Burn, Timothy C.
APPLICANT: Connors, Timothy D.
APPLICANT: Dackowski, William R.
APPLICANT: Van Raay, Terence J.
APPLICANT: Klinger, Katherine W.
TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
METHODS OF MAKING AND USING SAME
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESS:
ADDRESS: GENZYME CORPORATION
STREET: One Mountain Road
CITY: Framingham
STATE: Massachusetts

US-08-762-500-19

```
COUNTRY: United States of America
ZIP: 01701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/762,500
FILING DATE: 09-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/665,259
FILING DATE: 17-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCI/0596/10469
FILING DATE: 17-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dugan, Deborah A.
REGISTRATION NUMBER: 97,315
REFERENCE/DOCKET NUMBER: IG5-9.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508) 872-8400
TELEFAX: (508) 872-5415
INFORMATION FOR SEQ. ID NO: 19
SEQUENCE CHARACTERISTICS:
LENGTH: 6803 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-762-500-19

Query Match 1.18; Score 17, DB 3, Length 6803,
Best Local Similarity 100.0%; Pred. No. 69;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 398 GTGTCCAGGTCCCGAG 414
DB 4097 GTGTCCAGGTCCCGAG 4081

RESULT 28
US-08-222-617A-26
Sequence 26, Application US/08222617A
Patent No. 5882879
GENERAL INFORMATION:
APPLICANT: Veenstra, Annemarie E.
APPLICANT: Martin, Juan F.
APPLICANT: Garcia, Bruno D.
APPLICANT: Gutierrez, Santiago
APPLICANT: Barredo, Jose L.
APPLICANT: Von Boehren, Hans
APPLICANT: Palissa, Harriet
APPLICANT: Van Liempt, Henk
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/762,500
FILING DATE: 09-DEC-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
ATTORNEY/AGENT INFORMATION:
REFERENCE/DOCKET NUMBER: 97,157
INFORMATION FOR SEQ. ID NO: 26;
SEQUENCE CHARACTERISTICS:
LENGTH: 11444 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-222-617A-26

Query Match 1.18; Score 17, DB 3, Length 6803,
Best Local Similarity 100.0%; Pred. No. 69;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 398 GTGTCCAGGTCCCGAG 414
DB 4097 GTGTCCAGGTCCCGAG 4081

RESULT 28
US-08-222-617A-26
Sequence 26, Application US/08222617A
Patent No. 5882879
GENERAL INFORMATION:
APPLICANT: Veenstra, Annemarie E.
APPLICANT: Martin, Juan F.
APPLICANT: Garcia, Bruno D.
APPLICANT: Gutierrez, Santiago
APPLICANT: Barredo, Jose L.
APPLICANT: Von Boehren, Hans
APPLICANT: Palissa, Harriet
APPLICANT: Van Liempt, Henk
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/762,500
FILING DATE: 09-DEC-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
ATTORNEY/AGENT INFORMATION:
REFERENCE/DOCKET NUMBER: 97,157
INFORMATION FOR SEQ. ID NO: 1;
SEQUENCE CHARACTERISTICS:
LENGTH: 12364 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Penicillium chrysogenum
FEATURE:
NAME/KEY: CDS
LOCATION: 264..11600
```

```
APPLICATION NUMBER: US/08/222,617A
FILING DATE: 04-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
REFERENCE/DOCKET NUMBER: 97,157
INFORMATION FOR SEQ. ID NO: 26;
SEQUENCE CHARACTERISTICS:
LENGTH: 11444 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-222-617A-26

Query Match 1.18; Score 17, DB 3, Length 11;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 17; Conservative 0; Mismatches 0; Indels 0;

QY 668 ATGCTTAAAGAGAGAA 984
DB 7740 ATGCTGAGAGAGAGAA 7756

RESULT 29
US-08-222-617A-1
Sequence 1, Application US/08222617A
Patent No. 5882879
GENERAL INFORMATION:
APPLICANT: Veenstra, Annemarie E.
APPLICANT: Martin, Juan F.
APPLICANT: Garcia, Bruno D.
APPLICANT: Gutierrez, Santiago
APPLICANT: Barredo, Jose L.
APPLICANT: Von Boehren, Hans
APPLICANT: Palissa, Harriet
APPLICANT: Van Liempt, Henk
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/762,500
FILING DATE: 04-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
ATTORNEY/AGENT INFORMATION:
REFERENCE/DOCKET NUMBER: 97,157
INFORMATION FOR SEQ. ID NO: 1;
SEQUENCE CHARACTERISTICS:
LENGTH: 12364 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Penicillium chrysogenum
FEATURE:
NAME/KEY: CDS
LOCATION: 264..11600
```

```
OTHER INFORMATION: /function="enzyme"
OTHER INFORMATION: /product="AcV Synthetase"
US 08-222-617A-1
Query Match 13; Score 17; DB 2; Length 1254;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 968 ACTGCTGAGGACGAGAA 984
|||||
DB 7740 ACTGCTGAGGACGAGAA 7756

RESULT 30
US 09-754-250 3
Sequence 3; Application US/09754250
Patent No. 6376225
GENERAL INFORMATION:
APPLICANT: WEI, Ming-Hui et al
TITLE OF INVENTION: ISOLATED HUMAN PHOSPHODIESTERASE
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
FILE REFERENCE: C1001063
CURRENT APPLICATION NUMBER: US/09/754,250
CURRENT FILING DATE: 2001-01-05
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 111282
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(111282)
OTHER INFORMATION: n = A,T,C or G
US-09-754-250-3

Query Match 118; Score 17; DB 4; Length 111282;
Host Local Similarity 100.0%; Pred. No. 73;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 25 TGAATGAGGAGTCCTT 41
|||||
DB 3656 TGAATGAGGAGTCCTT 36582

RESULT 31
US-09-676-610H-24/c
Sequence 24; Application US/09676610H
Patent No. 6444465
GENERAL INFORMATION:
APPLICANT: C. Frank Bennett
APPLICANT: Jacqueline Wyatt
APPLICANT: Susan M. Preler
TITLE OF INVENTION: OLIGONUCLEOTIDE INHIBITION OF HEP-1 EXPRESSION
FILE REFERENCE: R15-0138
CURRENT APPLICATION NUMBER: US/09/676,610H
CURRENT FILING DATE: 2000-09-29
NUMBER OF SEQ ID NOS: 182
SEQ ID NO 24
LENGTH: 169998
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: exon
LOCATION: (1208)...(1472)
NAME/KEY: intron
LOCATION: (1473)...(124390)
NAME/KEY: exon
LOCATION: (124391)...(124544)
NAME/KEY: intron
LOCATION: (124545)...(125409)
NAME/KEY: exon
```

```
LOCATION: (125410)...(125595)
NAME/KEY: intron
LOCATION: (125596)...(128711)
NAME/KEY: exon
LOCATION: (128712)...(128848)
NAME/KEY: intron
LOCATION: (128849)...(133400)
NAME/KEY: exon
LOCATION: (133401)...(133469)
NAME/KEY: intron
LOCATION: (133470)...(134652)
NAME/KEY: exon
LOCATION: (134653)...(134773)
NAME/KEY: intron
LOCATION: (134774)...(136116)
NAME/KEY: exon
LOCATION: (136117)...(136261)
NAME/KEY: intron
LOCATION: (136262)...(137936)
NAME/KEY: exon
LOCATION: (137937)...(138053)
NAME/KEY: intron
LOCATION: (138054)...(138637)
NAME/KEY: exon
LOCATION: (138638)...(138766)
NAME/KEY: intron
LOCATION: (138767)...(138864)
NAME/KEY: exon
LOCATION: (138865)...(138940)
NAME/KEY: intron
LOCATION: (138941)...(139765)
NAME/KEY: exon
LOCATION: (139766)...(139860)
NAME/KEY: intron
LOCATION: (139861)...(142245)
NAME/KEY: exon
LOCATION: (142246)...(142445)
NAME/KEY: intron
LOCATION: (142446)...(143605)
NAME/KEY: exon
LOCATION: (143606)...(143738)
NAME/KEY: intron
LOCATION: (143739)...(145838)
NAME/KEY: exon
LOCATION: (145839)...(145931)
NAME/KEY: intron
LOCATION: (145932)...(147385)
NAME/KEY: exon
LOCATION: (147386)...(147544)
NAME/KEY: intron
LOCATION: (147545)...(153274)
NAME/KEY: exon
LOCATION: (153275)...(153321)
NAME/KEY: intron
LOCATION: (153322)...(155088)
NAME/KEY: exon
LOCATION: (155089)...(155231)
NAME/KEY: intron
LOCATION: (155232)...(156025)
NAME/KEY: exon
LOCATION: (156026)...(156151)
NAME/KEY: intron
LOCATION: (156152)...(156826)
NAME/KEY: exon
LOCATION: (156827)...(156928)
NAME/KEY: intron
LOCATION: (156929)...(163399)
NAME/KEY: exon
LOCATION: (163400)...(163586)
US-09-676-610H-24

Query Match 118; Score 17; DB 4; Length 169998;
Best Local Similarity 100.0%; Pred. No. 74;
```



```

ADDRESS: LAHIVE & COCKFIELD LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,507
FILING DATE: 15-May-1995
CLASSIFICATION:

PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/969,875
FILING DATE: 30-October-1992
ATTORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: 1M1-049DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ. ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 368 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 3..368
US-08-441-507-22

Query Match          1.0%; Score 16; DB 4; Length 368;
Best Local Similarity 100.0%; Prod. No. 2c+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 319 TCTTGCCATGCGCG 344
      |||
      433 TCTTGCCATGCGCG 318

RESULT 36
US-07-969-875A-22/C
Sequence 22, Application US/07/969,875A
Patent No. 6441157
GENERAL INFORMATION:
APPLICANT: Singh, Mohan Brij
APPLICANT: Smith, Penelope and
APPLICANT: Knox, Robert Bruce
TITLE OF INVENTION: Protein Allergens of the Species
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESS: Immunologic Pharmaceutical Corporation
STREET: 610 Lincoln Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/969,875A
FILING DATE: 30-October-1992
CLASSIFICATION: 435

```

```

ATTORNEY/AGENT INFORMATION:
NAME: Vanshoni, Darlene A.
REGISTRATION NUMBER: 36,279
REFERENCE/DOCKET NUMBER: 041.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 466-6000
TELEFAX: (617) 466-6010
INFORMATION FOR SEQ. ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 368 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 3..368
US-07-969-875A-22

Query Match          1.0%; Score 16; DB 4; Length 368;
Best Local Similarity 100.0%; Prod. No. 2c+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 TCTTGCCATGCGCG 344
      |||
      433 TCTTGCCATGCGCG 318

Db 333 TCTTGCCATGCGCG 318

RESULT 37
US-08-441-507-17/C
Sequence 17, Application US/08441507
Patent No. 6214358
GENERAL INFORMATION:
APPLICANT: Singh, Mohan Brij
APPLICANT: Smith, Penelope and
APPLICANT: Knox, Robert Bruce
TITLE OF INVENTION: Protein Allergens of the Species Cynodon
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESS: LAHIVE & COCKFIELD LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,507
FILING DATE: 15-May-1995
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/969,875
FILING DATE: 30-October-1992
ATTORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: 1M1-049DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ. ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 594 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-441-507-17

```

Query Match 1.08; Score 16; DB 4; Length 594;
Best Local Similarity 100.0%; Pred. No. 26-02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 419 TCTTGCCCATGGCGGC 334
|||||
DB 47 TCTTGCCCATGGCGGC 22

RESULT 48

US-07-969-875A-17/c

Sequence 17, Application US/07969875A

Patent No. 6441157

GENERAL INFORMATION:

APPLICANT: Singh, Mohan Hir;

APPLICANT: Smith, Penelope; and

APPLICANT: Knox, Robert Bruce

TITLE OF INVENTION: Protein Allergens of the Species

NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunologic Pharmaceutical Corporation

STREET: 610 Lincoln Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02154

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/969,875A

FILING DATE: 30-October-1992

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Vanstone, Darlene A.

REGISTRATION NUMBER: 35,279

REFERENCE/DOCKET NUMBER: 041.1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 466-6000

TELEFAX: (617) 466-6010

INFORMATION FOR SEQ. ID NO: 17:

SEQUENCE CHARACTERISTICS:

LENGTH: 594 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-07-969-875A-17

Query Match 1.08; Score 16; DB 4; Length 594;
Best Local Similarity 100.0%; Pred. No. 26-02;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 419 TCTTGCCCATGGCGGC 334
|||||
DB 37 TCTTGCCCATGGCGGC 22

RESULT 49

US-08-441-507-1/c

Sequence 1, Application US/08441507

Patent No. 6214358

GENERAL INFORMATION:

APPLICANT: Singh, Mohan Hir;

APPLICANT: Smith, Penelope; and

APPLICANT: Knox, Robert Bruce

TITLE OF INVENTION: Protein Allergens of the Species Cynodon

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: LATIVE & COCKFIELD LLP

STREET: 28 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/441,507

FILING DATE: 15-May-1995

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/969,875

FILING DATE: 30-October-1992

ATTORNEY/AGENT INFORMATION:

NAME: Mandragouras, Amy E.

REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: 1M1-0450V

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 742-4214

INFORMATION FOR SEQ. ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 662 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..435

US-08-441-507-1

Query Match 1.08; Score 16; DB 4; Length 662;
Best Local Similarity 100.0%; Pred. No. 2,10-02;
Matches 16; Conservative 0; Mismatches 0; Indels 0;

QY 319 TCTTGCCCATGGCGGC 334
|||||
DB 58 TCTTGCCCATGGCGGC 43

RESULT 40

US-07-969-875A-1/c

Sequence 1, Application US/07969875A

Patent No. 6441157

GENERAL INFORMATION:

APPLICANT: Singh, Mohan Hir;

APPLICANT: Smith, Penelope; and

APPLICANT: Knox, Robert Bruce

TITLE OF INVENTION: Protein Allergens of the Species

NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunologic Pharmaceutical Corporation

STREET: 610 Lincoln Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02154

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/969,875A

FILING DATE: 30-October-1992

```

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vanstone, Darlene A.
REGISTRATION NUMBER: 45,279
REFERENCE/DOCKET NUMBER: 041.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 466-6000
TELEFAX: (617) 466-6000
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 662 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..435
US-07-969-875A-1

```

```

Query Match: 1.0%; Score 16; DB 4; Length 662;
Best Local Similarity: 100.0%; Freq. No. 2.1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

CY 319 TCTTGCCATGCGCC 334
|||||
Fb 58 TCTTGCCATGCGCC 43

```

```

RESULT 41
US-09-288-143-11
Sequence 11, Application US/09288143
Patent No. 6433139
GENERAL INFORMATION:
APPLICANT: Brewer et al.
TITLE OF INVENTION: 53 Human Secreted Proteins
FILE REFERENCE: P2018P1
CURRENT APPLICATION NUMBER: US/99/288,143
CURRENT FILING DATE: 1999-04-08
EARLIER APPLICATION NUMBER: PCT/US98/21142
EARLIER FILING DATE: 1998-10-08
EARLIER APPLICATION NUMBER: 60/061,463
EARLIER FILING DATE: 1997-10-09
EARLIER APPLICATION NUMBER: 60/061,529
EARLIER FILING DATE: 1997-10-09
EARLIER APPLICATION NUMBER: 60/071,498
EARLIER FILING DATE: 1997-10-09
EARLIER APPLICATION NUMBER: 60/061,527
EARLIER FILING DATE: 1997-10-09
EARLIER APPLICATION NUMBER: 60/061,536
EARLIER FILING DATE: 1997-10-09
EARLIER APPLICATION NUMBER: 60/061,542
EARLIER FILING DATE: 1997-10-09
NUMBER OF SEQ ID NOS: 219
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 11
LENGTH: 680
TYPE: DNA
ORGANISM: Homo sapiens
US-09-288-143-11

```

```

Query Match: 1.0%; Score 16; DB 4; Length 680;
Best Local Similarity: 100.0%; Freq. No. 2.1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

CY 1527 TAACTGCAAAAAA 1542
|||||
Db 640 TAACTGCAAAAAA 655

```

```

RESULT 42
US-09-198-603C-26
Sequence 20, Application US/09198603C

```

```

Patent No. 6347194
GENERAL INFORMATION:
APPLICANT: TULLY, Raymond E.
APPLICANT: CALTAGIRONE, G. Thomas
APPLICANT: MOYER, Shawn S.
APPLICANT: KENNEDY, Michael J.
TITLE OF INVENTION: EXPRESSION OF MANNOSYL-BINDING PROTEIN IN METHYLOTROPHIC YEAST
FILE REFERENCE: A7290
CURRENT APPLICATION NUMBER: US/99/198,603C
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 26
LENGTH: 714
TYPE: DNA
ORGANISM: CHICKEN
US-09-198-603C-26

```

```

Query Match: 1.0%; Score 16; DB 4; Length 714;
Best Local Similarity: 100.0%; Freq. No. 2.1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

CY 1139 AATGAAATCTGCGG 1154
|||||
Db 628 AATGAAATCTGCGG 644

```

```

RESULT 43
US-08-441-507-20/C
Sequence 20, Application US/08441507
Patent No. 6214358
GENERAL INFORMATION:
APPLICANT: Smith, Mohan Biji;
APPLICANT: Smith, Penelope; and
APPLICANT: Knux, Robert Bruce
TITLE OF INVENTION: Protein Allergens of the Species Cynodon
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSER: LAHIVE & CROCKETT LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
METHOD TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,507
FILING DATE: 15-May-1995
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/969,875
FILING DATE: 30-October-1992
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: 1M1-0490V
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 759 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:

```


NAME/KEY: CDS
LOCATION: 1..738
FEATURE:
NAME/KEY: CDS
LOCATION: 742 759
US-08-441-507-20

Query Match 1.08; Score 16; DB 4; Length 759;
Best Local Similarity 100.0%; Pred. No. 2,1e-02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 TCTGGCCATGGCGC 334
|||||
DB 355 TCTGGCCATGGCGC 340

RESULT 44

US-07-969-875A-20/C
Sequence 20, Application US/07969875A
Patent No. 6441157

GENERAL INFORMATION:
APPLICANT: Singh, Mohan Brij;
APPLICANT: Smith, Penelope; and
APPLICANT: Knox, Robert Bruce
TITLE OF INVENTION: Protein Allergens of the Species
NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:
ADDRESS: Immunologic Pharmaceutical Corporation
STREET: 610 Lincoln Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02154

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/969,875A
FILING DATE: 30-October-1992
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Vanstone, Darlene A.;
REGISTRATION NUMBER: 35,279
REFERENCE/DOCKET NUMBER: 041.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 466-6000
TELEFAX: (617) 466-6010

INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 759 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA

FEATURE:
NAME/KEY: CDS
LOCATION: 1..738
NAME/KEY: CDS
LOCATION: 742..759
US-07-969-875A-20

Query Match 1.08; Score 16; DB 4; Length 759;
Best Local Similarity 100.0%; Pred. No. 2,1e-02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 TCTGGCCATGGCGC 334
|||||
DB 355 TCTGGCCATGGCGC 340

RESULT 45
US-08-441-507-3/C
Sequence 3, Application US/08441507
Patent No. 6214358

GENERAL INFORMATION:
APPLICANT: Singh, Mohan Brij;
APPLICANT: Smith, Penelope; and
APPLICANT: Knox, Robert Bruce
TITLE OF INVENTION: Protein Allergens of the Species Cyt
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESS: LAHIVE & COCKFIELD LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,507
FILING DATE: 15-May-1995
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/969,875
FILING DATE: 30-October-1992
ATTORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: 1M1-049DV

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 775 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA

FEATURE:
NAME/KEY: CDS
LOCATION: 1..600
US-08-441-507-3

Query Match 1.08; Score 16; DB 4; Length 775;
Best Local Similarity 100.0%; Pred. No. 2,1e-02;
Matches 16; Conservative 0; Mismatches 0; Indels 0;

QY 319 TCTGGCCATGGCGC 334
|||||
DB 217 TCTGGCCATGGCGC 202

Search Completed: December 7, 2002, 08:48:07
Job time : 649 secs

...

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2002, 13:53:37 Search time 2254 Seconds

(without alignments)
2335.197 Million cell updates/sec

Title: US-09-831-458A-12

Perfect score: 325

Sequence: 1 MSDSKPRVQGLGLCLGLH..

Scoring table:

OLIGO	60.0	Xgapex	60.0
Xgapop	60.0	Xgapex	60.0
Ygapop	60.0	Ygapex	60.0
Delop	6.0	Delop	7.0
	6.0	Delop	7.0

Searched: 16154066 seqs, 8097743376 residues

Word size: 5

Total number of hits satisfying chosen parameters: 2258837

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Command line parameters:

MODEL-frame-p2n.model -DBV-X1h
-O/cen2_1705PT01/seq81458/runtat_05122002_104421_169377/APP-query.fasta_1.519
-DB-EST-05MT-lastap -SUFFIX-011.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEX=0
-UNITS-bits -START=1 -END=1 -MATRIX=oligo -TRANS-human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=quality -THR MIN=5 -ALIGN=45 -MOPR=local -OUTPRMT=PT0
-NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-NO_XLPEX -NO_MMAP -LARGEQUERY -NEM_SCORES=0 -WAIT -CONVLOC -REV_TIMEOUT=120
-MAIN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YCAPOP=60 -YCAPEXT=60 -DELOP=6 -DPLEXT=7

Database:

EST: *
1: em_estba: *
2: em_esthum: *
3: em_estinu: *
4: em_estinu: *
5: em_estinu: *
6: em_estinu: *
7: em_estinu: *
8: em_estinu: *
9: qd_est1: *
10: qd_est2: *
11: qd_est3: *
12: qd_est3: *
13: qd_est4: *
14: qd_est5: *
15: em_estfun: *
16: em_estfun: *
17: qd_est1: *
18: qd_est1: *
19: em_gss_hum: *
20: em_gss_hum: *
21: em_gss_hum: *
22: em_gss_hum: *
23: em_gss_hum: *
24: em_gss_hum: *
25: em_gss_hum: *
26: em_gss_hum: *
27: em_gss_hum: *

Pred. No. is the number of results predicted by chance 1
score greater than or equal to the score of the result
and is derived by analysis of the total score distrib

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Accession	Species	Score
1	153	47.1	762	14	B1828875	Human	153.42
2	153	47.1	762	9	AL584658	Human	153.42
3	128	39.4	706	10	AV645755	Human	128.76
4	125	38.5	786	13	B1827955	Human	125.76
5	125	38.5	786	10	B1827955	Human	125.76
6	107	32.9	503	14	R98113	Human	107.76
7	97	29.8	792	9	AU140165	Human	97.76
8	60	18.5	419	14	H90968	Human	60.76
9	50	15.4	486	14	R80140	Human	50.76
10	47	14.5	833	9	AL570816	Human	47.76
11	45	13.8	359	10	AV645629	Human	45.76
12	39	12.0	411	9	AA448145	Human	39.76
13	19	5.8	453	14	R80141	Human	19.76
14	15	4.6	499	9	AA441250	Human	15.76
15	11	3.4	284	9	AL614060	Human	11.76
16	11	3.4	388	9	AA510960	Human	11.76
17	11	3.4	635	10	B18283702	Human	11.76
18	10	3.1	666	17	AG139337	Human	10.76
19	10	3.1	666	17	AG139337	Human	10.76
20	9	2.8	69	17	AC748251	Human	9.76
21	9	2.8	69	17	AC748251	Human	9.76
22	9	2.8	282	10	B1828875	Human	9.76
23	9	2.8	288	17	B1828875	Human	9.76
24	9	2.8	300	14	C57235	Human	9.76
25	9	2.8	386	10	AME34972	Human	9.76
26	9	2.8	461	12	B3554984	Human	9.76
27	9	2.8	482	10	AA452345	Human	9.76
28	9	2.8	489	9	AA388306	Human	9.76
29	9	2.8	591	12	B1828875	Human	9.76
30	9	2.8	607	17	B1828875	Human	9.76
31	9	2.8	610	13	B1481901	Human	9.76
32	9	2.8	614	10	B1828875	Human	9.76
33	9	2.8	633	10	B1828875	Human	9.76
34	9	2.8	719	14	B0744642	Human	9.76
35	9	2.8	810	17	CNS02X10	Human	9.76
36	9	2.8	811	17	CNS02X10	Human	9.76
37	9	2.8	879	14	B1757640	Human	9.76
38	9	2.8	897	17	CNS04Y80	Human	9.76
39	9	2.8	936	17	CNS03H1K	Human	9.76
40	9	2.8	977	17	CNS022K0	Human	9.76
41	9	2.8	1018	17	B17648	Human	9.76
42	9	2.8	1101	17	CNS00385	Human	9.76
43	9	2.8	3108	11	AF126916	Human	9.76
44	8	2.5	164	13	B1828875	Human	8.76
45	8	2.5	171	14	B0744642	Human	8.76

ALIGNMENTS

RESULT 1
B1828875
LOCUS
DEFINITION
603075032F1 NIH_M01_119 Homo sapiens cDNA clone
B1828875
ACCESSION
B1828875.1 GI:15949425
VERSION
B1828875.1
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata
Mammalia; Eutheria; Primates; Catarrhini; Hominii
1 (bases 1 to 769)
REFERENCE
NIH-M01 http://www.ncbi.nlm.nih.gov/
AUTHORS
National Institutes of Health, Mammalian Gene Cl


```

DB 429 GAGTACGCTGAGTTCGCTAGAGAAATCTCAATCTTATAGATCTATAGATCTTAA 488
QY 146 ACGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 155
DB 489 GAGTACGCTGAGTTCGCTAGAGAAATCTCAATCTTATAGATCTTAAATCTTAAATCTTAA 548
QY 156 GATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 168
DB 549 GAGTACGCTGAGTTCGCTAGAGAAATCTCAATCTTATAGATCTTAAATCTTAAATCTTAA 597

RESULT 3
LOCUS AV645755 706 bp mRNA linear EST 15 JAN 2002
DEFINITION AV645755 GIC Homo sapiens cDNA clone GICAD05 Y, mRNA sequence.
ACCESSION AV645755
VERSION AV645755.1 GI:9866769
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 706)
AUTHORS Xu, X., Huang, J., Xu, Z., Qian, B., Zhu, Z., Yan, Q., Cai, T., Zhang, X.,
Xiao, H., Qu, J., Liu, F., Huang, Q., Cheng, Z., Li, N., Du, T., Hu, W.,
Shen, K., Lu, G., Fu, G., Zhong, M., Xu, S., Gu, W., Huang, W., Zhao, X.,
Hu, G., Gu, J., Chen, Z., and Han, Z.
TITLE Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (25), 15084-15094 (2001)
21625106
CONTACT: Zeguana Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhanjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex 45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
Source
Location/Qualifiers
1..706
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GICAD05"
/clone_lib="GIC"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 168 a 198 c 176 g 162 t 2 others
ORIGIN
XhoI"

Alignment Scores:
Pred. No.: 1 98e-122 length: 706
Score: 128.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 39.38% Indels: 0
DB: 10 gaps: 0

US-09-831-458a-12 (1-325) x AV645755 (1-706)
QY 188 GATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 207
DB 145 GAGTACGCTGAGTTCGCTAGAGAAATCTCAATCTTATAGATCTTAAATCTTAAATCTTAA 204
QY 208 MetStransSrrgIntraGASrTrpHisASrSeiValThrAlaGscIncluvAlArgAla 227
DB 205 ATGTCTACTGCTAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 264
QY 228 GATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 247

```

```

DB 205 GAGTACGCTGAGTTCGCTAGAGAAATCTCAATCTTATAGATCTTAAATCTTAAATCTTAA 204
QY 248 GATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 257
DB 325 GATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 344
QY 268 ValAspGlySerProLeuSerTrpGluPheGluAlaGlyTrpPAsuSerGly 287
DB 385 GATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 404
QY 288 AsnSerGlyAsnGluAspGlyAlaGluPheSerGlySerGlyTrpPAsuSerGly 307
DB 445 AATATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 464
QY 308 AspValAspAsnTrpTrpLeuGly 315
DB 505 GAGTACGCTGAGTTCGCTAGAGAAATCTCAATCTTATAGATCTTAAATCTTAAATCTTAA 524

RESULT 4
LOCUS B1827055 786 bp mRNA linear
DEFINITION B1827055 NIH-MGC_119 Homo sapiens cDNA clone
mRNA sequence.
ACCESSION B1827055
VERSION B1827055.1 GI:15948605
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea;
REFERENCE 1 (bases 1 to 786)
AUTHORS NIH-MGC http://www.ncbi.nlm.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: greg@nigms.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium
DNA Sequencing by: Incyte Genomes, Inc.
cDNA distribution: MGC clone distribution info:
http://image.llnl.gov
Plate: L14M11416 row: d column: 08
High quality sequence stop: 776.

FEATURES
Source
Location/Qualifiers
1..786
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:516747"
/clone_lib="NIH-MGC_119"
/tissue_type="medulla"
/lab_host="pBluscript"
/note="Organ: brain; Vector: pCMV-Sport;
Site_2: EcoRV (destroyed); RNA source:
anonymous male age 27. Library is
directly cloned, cloned (EcoRV site is
destroyed). Average insert size 1.3 kb,
0.9-3 kb. Library is normalized and
full-length clones and was constructed
(invitrogen). Research Genetics Trace
this is a NIH-MGC library."
BASE COUNT 212 a 193 c 241 g 140 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 3e-119 length: 786
Score: 125.00 Matches: 125
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 38.46% Indels: 0
DB: 13 gaps: 0

```

RESULT 10	
AL570816/c	833 bp MPNA linear EST 16-FEB-2001
LOCUS	
DEFINITION	AL570816 LTI_NFL006_PL2 Homo sapiens cDNA clone GSDDI02YH14.3
ACCESSION	AL570816
VERSION	AL570816.1 GI:12927495
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Hemo-
TITLE	1 (bases 1 to 833)
JOURNAL	Ll,h,r , Gruber,C , Jasson,J and Delays,D,
COMMENT	Full-length cDNA libraries and normalization Unpublished (2001) Contact : Genoscope Genoscope - Centre National de Sequenage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr. Location/Voualliers L . 833 /organism="Homo sapiens" /db.xref="taxon:9606" /clone="GSDDI02YH14" /clone_lib="LTI_NFL006_PL2" /issue_type="placenta" /note="Vector: pCMVSPORT 6, Site-1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9600 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 510 8371 Email : fliang@life tech.com URL : http://fulllength.invitrogen.com"
BASE COUNT	213 a 193 c 219 g 184 t 4 others
ORIGIN	
Alignment Scores:	
Pred. No.:	4,3e-38 Length: 833
Score:	47.00 Matches: 47
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	14.46% Indels: 0
DB:	9 Gaps: 0
US-09-B-I-458A-12 (1-325) x AL570816 (1 833)	
QY	275 ProSerPheGlnAlaGTYTTTPASGSeSgLYInPrAsnAsnSeGtGYASugUASpCYs 294
Dd	CCCCCCTTCATGGAGTCACTGCATCGAACATCCAAATAATGAGGAATCTT 0+4
QY	295 ALaGluPheSerGlySerGlyTTPASnASPASnArGYsAspValASPAnTYTTrPlJe 314
Dd	GGGGAATTACTGTGACTGTGGTAACACCAATGATGATGAGTGACAAATTCAGATC 604
QY	315 CysTlysProAlaAlaCYs 321
Dd	TTTTTTTTTTTTTTTTTTT 583
Db	603 TGC AAAAAGCCCGCACGCCGC 583
RESULT 11	
AV645629	359 bp MPNA linear EST 15-JAN 2002
LOCUS	
DEFINITION	AV645629 GLC Homo sapiens cDNA clone GICCD12.3', mRNA sequence.
ACCESSION	AV645629
VERSION	AV645629.1 GI:9866643
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	

REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	COMMENT
1	(bases 1 to 359) Xu, X., Huang, J., Xu, Z., Qian, H., Zhu, Z., Yan, G., Xiao, H., Gu, Z., Lin, F., Huang, G., Cheng, Z., Li, N. Shen, K., Lu, G., Fu, G., Zhong, M., Xu, S., Gu, W., Hu, G., Gu, J., Chen, Z., and Han, Z.	Insight into hepatocellular carcinoma pathogenesis at the level of gene expression profiles of hepatoma with those of corresponding noncancerous liver Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15092 (2001)			
	Contact: Zengqiang Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Shanghai Hi-Tech Park 201204 P. R. China Tel.: 86-21-50801919 ext. 457 Fax: 86-21-50801922 Email: hanzq@hgc.sh.cn				
	This clone is available at either in Shanghai. Location/Qualifiers 1..359 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="GICACB12" /clone_lib "SICF" /tissue="lymphocytes"corresponding non cancerous /dev_stage "Adult" /lab_host "SOLR" /note="Vector: pBluescript SK(-); Site: 1 XhoI"				
	BASE COUNT 85 a 111 c 91 g 72 t				
	ORIGIN				
	Alignment Scores:				
	Prepared. No.:	1.72e+46	Length:	359	
	Score:	45.00	Matches:	45	
	Percent Similarity:	100.00%	Conservative:	0	
	Best Local Similarity:	100.00%	Mismatches:	0	
	Query Match:	13.85%	Indels:	0	
	DB:	10	gaps:	0	
	US-09-831-458A-12 (1-325) x NW45629 (1-459)				
Q7	188 GI:474910CysARGHISCSYPTOLYSASPIRPHRPHAPGNGIYAS				
Db	145 GAACGCGCTGTGGCGGCACCTGTCCAGAGATGGACATTTCTTCCAGGAAAT				
QY	208 MetSerAsnSerGlnIArgAsnTrpHisAspSerValThrAlaCysGlnGln				
Db	205 ATGTCTAACCTCCACCGGAAATTCGTACGACTGGCTGCACCTGCCAGAAAT				
QY	228 GTCLeuValIle 232				
Db	265 CAGTCTGCTGATATC 279				
RESULT 12					
LOCUS	AA448145	411 bp	mRNA	1000	
DEFINITION	2882912.ct Scores_test.cts_MHT Homo sapiens cDNA				
ACCESSION	AA448145				
VERSION	AA448145.1				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Exaiyola, Metacay, Charata, Caramata, Veripont Mammalia: Eutheria, Primates, Catarrhini, Homini: 1 (bases 1 to 411) Hillier, L., Allen, M., Bowles, L., Bonquene, T., Mc Kusaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., P Schellenberg, K., Stepien, M., Tan, F., Theisind T., Waterston, R., and Wilson, R. WashU-Merck EST Project 1997				
TITLE	JOURNAL				
	Unpublished (1997)				

Tel: 314 286 1800

IMAGE Consortium (infoimage.jhu.edu) for further information.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;
 1 (bases 1 to 381)
 REFERENCE Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman,
 M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
 Rinkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisan, F., Waterston,
 R., Williamson, A., Wohlmann, P., and Wilson, R.
 The WashU-Merck EST Project
 TITLE Unpublished (1995)
 JOURNAL Contact: Wilson, R.
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8001, St. Louis, MO 63110
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Insert Size: 712
 High quality sequence stops: 311
 Source: IMAGE Consortium, LIND
 This clone is available royalty-free through LIND; contact the
 IMAGE Consortium (info@image.lind.gov) for further information.
 Insert Length: 712 Std Error: 0.90
 Seq primer: M13RPI
 High quality sequence stops: 331.
 Location/Qualifiers
 1..381
 /organism="Homo sapiens"
 /db_xref="GDB:3790305"
 /db_xref="taxon:9606"
 /clone="IMAGE:241372"
 /clone_lib="Soares fetal liver spleen INFIS"
 /sex="male"
 /dev_stage="20 week post conception fetus"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Organ: Liver and Spleen; Vector: pT73b (Pharmacia)
 with a modified polylinker; Site_1: Pac I; Site_2: Eco RI,
 1st strand cDNA was primed with a Pac I - oligo(dT) primer
 15' AACTGCAAGATTATTAATGATCTTTTCTTTTCTTTT 3';
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Pac I and cloned into the Pac I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization. Library
 constructed by Benito Soares and M. Fatima Bonaldo."
 BASE COUNT 68 a 97 c 88 g 125 t 3 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 4.7 length: 381
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Host Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.08% Indels: 0
 gaps: 0
 DB: 14
 US-09-831-458a-12 (1-325) x H90360 (1-381)
 QY 316 LysLysProAlaAlaCysPheArgAspGlu 325
 |||
 Db 2 AAAAAGCCGCGAGCGCTGCTTCAGACGAA 31
 RESULT 19
 AC141937/c AG143937 666 bp DNA linear GSS:08-JAN-2002
 LOCUS DEFINITION Pan troglodytes RNA, clone: RP43-004M13.17, genomic survey
 sequence.
 AC143937
 VERSION AG143937.1 GI:16673615
 KEYWORDS GSS.
 SOURCE Pan troglodytes male lymphocytes DNA, clone: lib:RP43-43 Chimpanzee
 Male BAC Library clone:RP43-004M13.17.
 Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.

REFERENCE 1
 AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, J. D.,
 Toki, Y., Watanabe, H., and Sakaki, Y.
 TITLE BAC end sequences of Library P01-43
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 666)
 AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, J. D.,
 Toki, Y., Watanabe, H., and Sakaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (02 AUG 2001) Asao Fujiyama, The Institute
 and Chemical Research (RIKEN), Genomic Sciences
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa,
 (E-mail: fujiyama@riken.go.jp, Tel: 81-45-503-9111,
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)
 Clones are derived from the chimpanzee BAC library
 end was generated during the RFLP process and may
 of clone tracking errors.
 COMMENT
 PRIMER Sequencing: 17
 LIBRARY
 Vector : pBAVex4
 R.Site 1 : EcoRI
 R.Site 2 : EcoRI.
 Location/Qualifiers
 1..666
 /organism="Pan troglodytes"
 /db_xref="taxon:9598"
 /clone="RP43-004M13.17"
 /sex="male"
 /cell_type="lymphocytes"
 /clone_lib="RP43-43 Chimpanzee Male BAC
 GSS."
 BASE COUNT 121 a 270 c 147 g 126 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 9.75 length: 666
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Host Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.08% Indels: 0
 DB: 17
 gaps: 0
 US-09-831-458a-12 (1-325) x AG143937 (1-666)
 QY 270 GlycerProLeuSerProSerPheGlnArg 279
 |||
 Db 320 GATTCCTCTCTGCGCCGATTCGAGCG 291
 RESULT 20
 A2782451 649 bp DNA linear GSS:07-FEB-2001
 LOCUS DEFINITION 2M002211R Mouse 1993 Plac13 ungGEM library Mus
 clone: U002-2M002-2118 R, RNA sequence.
 ACCESSION A2782451
 VERSION A2782451.1 GI:12916186
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
 Mammalia; Eutheria; Rodentia; Soturognathi; Mur
 1 (bases 1 to 69)
 REFERENCE Dunn, D., Aoyagi, A., Bartol, M., Brown, T., Jones,
 Islam, H., Longacre, S., Mahmoud, M., Meegen, E., P.
 M., Rose, M., Rose, E., Strick, F., Tinney, A.,
 and Wright, D., Weiss, R.
 TITLE Mouse whole genome scaffolding with paired end
 plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 Km 308, Biomedical Polymers Research Bldg., 20
 84112, USA

Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watsong.wustl.edu
 Library constructed by life technologies. DNA Sequencing by:
 Washington University Genome Sequencing Center
 Clone distribution: Xenopus clones from this library are available
 through the I.M.A.G.E Consortium/LLNL at: info@image.llnl.gov
 Seq primer: -40bp from ribco
 High quality sequence stop: 439
 location/qualifiers

FEATURES

source

1. 461
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone_image="408871"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: heart; Vector: pCMV-Sport6; Site_1: NotI;
 Site_2: SalI. Cloned unidirectionally. Primer oligo dt.
 Average insert size 1.6 kb. Constructed by life
 technologies."
 BASE COUNT 141 a 89 c 81 g 150 t
 ORIGIN

Alignment Scores:

Pred. No.: 66.1 Length: 461
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.77% Indels: 0
 DB: 12 Gaps: 0

US-09-831-458a-12 (1-325) x BG554984 (1-461)

UY 217 ASPserValThrAlaGlycInclGluVal 225

Db 19 GATTGAGTTGACGCTGCCAGGAAGTA 45

RESULT 27

AA452345/C

LOCUS

DEFINITION

AA452345 482 bp mRNA linear EST 17-FEB-2000

IMAGE:3068165 3, mRNA sequence.

AM452345

AM452345.1 GI:6993121

EST.

SOURCE

ORGANISM

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCI-CCAP http://www.ncbi.nlm.nih.gov/nciccap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: rgs@pds.nhlbi.nih.gov

Oligo-dt track not found. Not 1 site shown in beginning of sequence

is likely internal to the message. cDNA library preparation: M.B.

Scores lab clone distribution: NCI-CCAP clone distribution

information can be found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/bdrrp/image/image.html

Seq primer: M13 Forward

POLYA-No.

Location/Qualifiers

1. 482

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_image="3068165"

/clone_lib="NCI-CCAP_Subs"

/lab_host="DH10B (life technologies)"

/note="Vector: p773d-pap (Pharmacia) with a modified

polylinker. Site_1: Not I; Site_2: Eco RI; NCI-CCAP_Subs

is a subtracted library derived from NCI-CCAP_Sub4. The

NCI-CCAP_Subs library had 3 million pooled
 single-stranded DNA preparation of NCI-CCAP
 as a tracer in a subtractive hybridization
 experiment. The IMAGE pool (NCI-CCAP_Kid
 3334-3347, 6682-6685, 6798-6803 (IMAGE)
 1323376-1323915, 1356008-1356675, 1506755-
 NCI-CCAP_Kid5 pool 1 LLAM 3338-3342, 4742-
 (IMAGE Clones: 1323912-1325831, 1471866-
 1492104 1492255); NCI-CCAP_Lib pool 1 LL
 3851-3854 (IMAGE Clones: 1414920-1417555
); NCI-CCAP-NC4 pool 1 LLAM 3164-3167,
 3733-3735 (IMAGE Clones: 1257096-125866-
 1475592-1476747); NCI-CCAP_Pred pool 1
 2758-2759, 4062-4068 (IMAGE Clones: 98-
 1101192-1101459, 1217928-1220615); NC
 LLAM 2644-2652, 2871-2872 (IMAGE Clones:
 1144584-1145351). (10% of the driver p
 pool of 3,610 unpaired clones from NCI-CCAP
 Clones: 2708616-2710535) and NCI-CCAP
 Clones: 2710536-2712455) (10% of the
), plus a pool of 11,136 clones from NCI-CCAP
 Clones: 2712456-2723591) (10% of the dr
 plus a pool of 5,472 clones from NCI-CCAP
 Clones: 2723592-2728969) (70% of the dr
 Subtraction was performed as previously
 , Lennon & Soares (1990): Normalization
 Two Approaches to Facilitate Gene Disc
 Research 6, 791-806.
 TAG-LIB: NCI-CCAP-NC4
 TAG-TISSUE: sperm cell
 TAG_SEQ: AATATC

BASE COUNT

91 a 135 c 157 g 98 t 1 other

ORIGIN

Alignment Scores:

Pred. No.: 70.1 Length: 482
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.77% Indels: 0
 DB: 10 Gaps: 0

US-09-831-458a-12 (1-325) x AM452345 (1-482)

UY 34 AlaGlyValLeuValAlaIleGluVal 42

Db 425 GCTGGAGTCTTGCTGCCAIC116510 999

RESULT 28

AA388306/C

LOCUS

DEFINITION

AA388306 499 bp mRNA linear EST

IMAGE:775542 5, mRNA sequence.

AM452345

AM452345.1 GI:2011722

EST.

SOURCE

ORGANISM

house mouse.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Squirrelia; Mus;

1 (bases 1 to 499)

Marra, M., Hillier, L., Allen, M., Bowles, M., D. G.

Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J.

Schellenberg, K., Strohman, J., Tan, F., Underwood, J.

Theltinger, B., Wylie, J., Lennon, G., Soares, R., W.

Waterston, R.

The WashU-HMI Mouse EST Project

Unpublished (1996)

Contact: Maria M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO

Tel: 314 286 1800


```

seedling"
/lab_host="DH10B"
/Note="Vector: pBluescript II SK+, Site1: EcoRI, Site2:
XhoI. The cDNA library was constructed from mRNA isolated
from degenerating ocytodes of 2 week old seedlings from
p1468916. Complementary DNA was synthesized from mRNA
using a primer consisting of a poly(dT) sequence with a
XhoI restriction site. EcoRI adapters were ligated to the
blunt-ended cDNA fragments followed by XhoI digestion.
The cDNA fragments were directionally cloned into the
EcoRI XhoI restriction site of the pBluescript vector. The
ligated cDNA fragments were transformed into DH10B host
cells (GibcoBRL). This library was constructed in the
laboratory of Dr. Randy Shoemaker."

BASE COUNT      111 a      179 c      160 g      157 t
ORIGIN

Alignment Scores:
Pred. No.:      94.7      Length:      607
Score:          9.00      Matches:      9
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      2.77%      Indels:      0
DB:              12      Gaps:      0

US-09-831-458a-12 (1-325) x BG041052 (1-607)

QY      239 AsnPhelcuglnleuglntrseratg 247
Db      70 AACCTCCTCAGCTTCAACCTTCCG 96
|||||
RESULT 41
LOCUS      B1481901      610 bp      mRNA      linear      EST 28-AUG-2001
DEFINITION      B664518.5 prime RE Drosophila melanogaster normalized Embryo p14c-1
Drosophila melanogaster cDNA clone B64518 5, mRNA sequence.
ACCESSION      B1481901
VERSION      B1481901.1 GI:15321110
KEYWORDS      EST
SOURCE      fruit fly.
ORGANISM      Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pleiyota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      Stapleton, M., Brockstein, P., Hong, L., Tyler, D., Bernat, B., Carlson
J., Champe, M., Chavez, C., Dorsett, V., Farrant, D., Frisze, E., George
K., Gonzalez, M., Guarino, H., Harris, N., Li, P., Li, G., Mista, S.,
Mungall, C. J., Nuppon, T., Parlebo, T., Paragas, V., Park, S.,
Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Celisner, S. and Rubin
G. M.
B66P/HMI RE Drosophila EST project
Unpublished (2001)
JOURNAL
COMMENT      Contact: Stapleton, M.
B66P
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_estafruitfly_berkeley.edu
hit genomic AEU03469; arm: 4L1608441.9111671
estimated-cyto: 6128-61E1: 05/16/2001
Plate: RF 645 row: B column: 5
High quality sequence stop: 540.
location/Qualifiers
1..610
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="B64518"
/clone_lib="RE Drosophila melanogaster normalized Embryo
p14c-1"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DH5-alpha Tona"

```

```

/Note="Organ: embryo; Vector: p14c1 Site1:
BamHI. Library was kindly generated by Dr.
the RIKEN. The library was normalized and
Cre recombinase. Plasmid cDNA library."

BASE COUNT      202 a      117 c      146 g      145 t
ORIGIN

Alignment Scores:
Pred. No.:      95.3      Length:      610
Score:          9.00      Matches:      9
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      2.77%      Indels:      0
DB:              13      Gaps:      0

US-09-831-458a-12 (1-325) x B1481901 (1-610)

QY      272 ProLcuserProscorPheclArqFyr 280
Db      260 CCACCTTCGCCCTCTCCACAGCTAT 284
|||||
RESULT 32
LOCUS      B637495      614 bp      mRNA      linear
DEFINITION      B637495 RIKEN full-length enriched, adult male
musculus cDNA clone A50080M21 5', mRNA sequence.
ACCESSION      B637495
VERSION      B637495.1 GI:16473240
KEYWORDS      EST.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata
Mammalia; Eutheria; Rodentia; Sciurognathi; Murinae
1 (bases 1 to 614)
REFERENCE      Arakawa, T., Carninci, P., Furuta, S., Furuo, M., Hara,
A., Hiramoto, K., Hori, P., Ishii, Y., Ito, M., Kawai,
M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, E.,
Okazaki, Y., Okubo, T., Saito, P., Sakai, T., Sakai, Y.,
Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Taniguchi,
M., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESis (Arakawa, T., et al. 2001)
Unpublished (2001)
JOURNAL
COMMENT      Contact: Yoshinori Hayashizaki
Laboratory for Genome Exploration Research Group
Science Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research
1-7-22 Suehiro cho, Tsurumi Ky, Yokohama, Kanagawa
Tel.: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome-res.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y.,
M., Kondo, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subfraction of cap-trapper sets
prepare full-length cDNA libraries for rapid dis-
covery. Genome Res. 10 (10): 1617-1630 (2000)
Waghi, K., Fujitake, S., Inoue, K., Ito, Y., Ito, Y.,
Watanabe, M., Yoshida, T., Ishikawa, T., Ozawa, K.,
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y.,
Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system
sequencing pipeline with 484 multiplexed cap-trapper
10 (11): 1757-1771 (2000)
Kondo, H., Fukunishi, Y., Shibata, K., Itoh, M., Goto,
Y., and Hayashizaki, Y.
Computer-based methods for the mouse full-length
cDNA library. real-time sequence clustering for
nonredundant cDNA library. Genome Res. 11 (2):
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H.,
K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata,
Hayashizaki, Y.
Computational Analysis of Full-Length Mouse cDNAs

```


TITLE Human gene number estimate provided by genome wide analysis using
 Tetraodon nigroviridis DNA sequence
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 811)
 AUTHORS Forest-Collins,H., Billon,C., Bastiva,C., Fitzames,C., Fischer,C.,
 Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
 Weissenbach,J.
 TITLE Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 811)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (12-APR-2000)
 COMMENT This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the Tetraodon nigroviridis
 genome. For more information, please take a look at
 http://www.genoscope.cns.fr/Tetraodon.
 FEATURES
 source location/Qualifiers
 1..811
 /organism="Tetraodon nigroviridis"
 /db_xref="taxon:99883"
 /clone_11b="g"
 /note="Genoscope sequence ID : COAG178D005LP1-end : 77"
 BASE COUNT 213 a 193 c 169 g 206 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 138 Length: 811
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservat: 9
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.77% Indels: 0
 Gaps: 0
 US-09-831-458a-12 (1-325) x CNS02XKQ (1-811)
 CY 22 AAAAAAAAAAAAAAAAAAAAAA
 DB 764 GCGTCCTGCTGCAGCTCTCTTCTT 738
 RESULT 37 879 bp mRNA linear EST 25-SEP-2001
 ID US B1757640/1
 DEFINITION 603027936F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE-5198265 5',
 mRNA sequence.
 ACCESSION B1757640
 VERSION B1757640
 KEYWORDS EST.
 SOURCE human
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 TITLE NIH-MGC http://mgi.mcg.mil.nih.gov/
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: rgs@bimail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I M A G E Consortium (LMNI)
 Clone distribution: MGC clone distribution information can be
 found through the I M A G E Consortium/LMNI at:
 http://image.llnl.gov
 plate: LLAM11496 row: h column: 10
 High quality sequence stop: 410.
 Location/Qualifiers
 1..879
 /organism="Homo sapiens"

/db_xref="taxon:9606"
 /clone_11b="NIH_MGC_114"
 /clone_11b="NIH_MGC_114"
 /lab_host="DH10B"
 /note="Human brain: Vector: pCMV-Sport6; Site_1: NotI;
 Site_2: EcoRV (destroyed); DNA source: anonymous pool of 6
 male brains, age range 23-27 yrs. Library is oligo dt
 primed and directionally cloned (Kovov site is destroyed
 upon cloning). Average insert size 1.5 kb. Insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code: 019. Note:
 this is a NIH-MGC Library."
 BASE COUNT 238 a 227 c 253 g 161 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 154 Length: 879
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservat: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.77% Indels: 0
 Gaps: 0
 US-09-831-458a-12 (1-325) x B1757640 (1-879)
 CY 16 GTCCTGCTGCAGCTCTCTTCTT 24
 DB 137 GCGTCCTGCTGCAGCTCTCTTCTT 111
 RESULT 38 897 bp DNA linear CDS 26 JUL 2000
 ID US B1757640/1
 DEFINITION Tetraodon nigroviridis genome survey sequence T3 end of clone
 region of library A131334.1 in tetraodon nigroviridis, genome survey
 sequence.
 ACCESSION A131334.1
 VERSION A131334.1 GI:3546227
 KEYWORDS GSS: genome survey sequence.
 SOURCE Tetraodon nigroviridis.
 ORGANISM Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodon.
 TITLE 1 (bases 1 to 897)
 JOURNAL Forest-Collins,H., Billon,C., Bastiva,C., Bouneau,L., Fischer,C.,
 Bernot,A., Fitzames,C., Winkler,P., Brottier,P., Quetier,F.,
 Saurin,W. and Weissenbach,J.
 COMMENT Estimate of human gene number provided by genome-wide analysis
 using Tetraodon nigroviridis DNA sequence
 REFERENCE NIH_MGC_114
 MEDLINE 20296634
 PUBMED 10845645
 JOURNAL 2 (bases 1 to 897)
 MEDLINE 20359837
 PUBMED 10890144
 REFERENCE 3 (bases 1 to 897)
 JOURNAL Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (12-APR-2000)
 COMMENT This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the Tetraodon nigroviridis
 genome. For more information, please take a look at
 http://www.genoscope.cns.fr/Tetraodon.
 FEATURES
 source location/Qualifiers
 1..897

RESULT 41
B12648 1018 bp DNA linear GSS 14 MAY 1997
DEFINITION F26F4-17.1 rat Arabidopsis thaliana genomic clone F26F4. DNA sequence.
ACCESSION B12648
VERSION B12648
KEYWORDS GSS
SOURCE thale cross.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Etrypophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis 1 (bases 1 to 1018)
REFERENCE Fennel, J., Dewar, K., Buehler, F., Kim, C., Li, Y., Shinn, P., Sun, H. and Ecker, J.
AUTHORS BAC End Sequences at ATCC
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: F26F4-Sp6, F26F4-T7, F26F4-SF6.1
Contact: Ecker, J.
Arabidopsis Thaliana Genome Center
University of Pennsylvania
Dept. of Biology, University of Pennsylvania, Philadelphia, PA 19104
Tel: 215-898-9384
Fax: 215-898-8780
Email: jecker@atgenome.bio.upenn.edu
Seq primer: T7
Class: BAC ends
High quality sequence start: 288
High quality sequence stop: 355.
Location/Qualifiers
1..1018
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="F26F4"
/sex="hermaphrodite"
/note="Vector: BelosAGC1; Site_1: EcoRI; Site_2: EcoRI; Produced by Thomas Altmann"
BASE COUNT 338 a 192 c 163 g 298 t 27 others
ORIGIN

Alignment Scores:
Pred. No.: 186 Length: 1018
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.778 Indels: 0
DB: 17 Gaps: 0

US-09-831-458a-12 (1-325) x B12648 (1-1018)
Q7 211 SerClaraGASrTriplussAsserVal 219
|||||
Db 614 TCGCAAGAAACCTGCTATCCTCTCTT 640

RESULT 42
CNS00385 1101 bp DNA linear GSS 04-JUN-1999
DEFINITION BAC008624 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION CNS00385
VERSION AL063897.1 GI:4941754
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Insecta; Ectoptera; Neoptera; Endopterygota; Diptera; Brachycera; Muscivora; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)

AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91106 Evry Evry-FRANCE (E-mail: seq@genoscope.cns.fr - Web: www.genoscope.cns.fr)
COMMENT Determination of this BAC end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org/TheBDGP/Drosophila_melanogaster_BAC_library was prepared by Katsuyo Osucawa and Aaron Mammoss in Peter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BAC08624"
/clone_lib="RPCI-98"
/note="end : 17"
BASE COUNT 223 a 238 c 232 g 374 t 34 others
ORIGIN

Alignment Scores:
Pred. No.: 206 Length: 1101
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.778 Indels: 0
DB: 17 Gaps: 0

US-09-831-458a-12 (1-425) x CNS00385 (1-1101)
Q7 272 ProteoSerProGrrPhoGicArTyr 280
|||||
Db 191 GCACTTGACCTGCTCCAGGCTTAT 217

RESULT 43
AF326918 3108 bp mRNA linear HTG 13-JUL-2001
DEFINITION Homo sapiens metalloproteinase-disintegrin meltrin beta (FMSG34)
ACCESSION AF326918
VERSION AF326918.1 GI:12276179
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Cladacea; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE Wang, Y.-G. and Gong, J., Identification of FMSG34, a novel human gene encoding for metalloproteinase-disintegrin meltrin beta unpublished 2 (bases 1 to 3108)
AUTHORS Wang, Y.-G.
TITLE Direct Submission
JOURNAL Submitted (07-DEC-2000) Beijing Fengkesheng Function Gene Technology Ltd., 4 Tong Tiao Lu Chang Street, Xuanwu District, Beijing 100050, P.R. China
REFERENCE 1 (bases 1 to 3108)
Location/Qualifiers
1..3108
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
1..3108
gene

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Comugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 7, 2002, 13:53:27 ; Search time 2755 seconds
(without alignments)

3431.180 Million cell updates/sec

Title: US-09-831-458a-12

Percent score: 325
Sequence: 1 MSRSKPPHVOQLGLGCLGH.....KQDVNWICKKPAACEFHE 325

Scoring table:
OLIGO
Xgapop 60.0 ; Xgapext 60.0
Ygapop 60.0 ; Ygapext 60.0
Kgapop 6.0 ; Kgapext 7.0
Delop 6.0 ; Deltext 7.0

Searched: 2054640 seqs, 14551402878 residues

Word size: 5

Total number of hits satisfying chosen parameters: 581759

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL-frame-p2n.model -DEV=xlh
-Q/cgnt2.1/USPTO.spool/US09831458/runat_05122002_103421_16305/36f-4seqv.fasta.1.513
-DB-GenBml -QFMT=fastap -SUFFIX=oli.rge -MINMATCH=0 -I=loopcl=0 -I=loopext=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=humand0.cdi -LIST=45
-DOCALLION=200 -THR_SCORE=quality -THR_MIN=5 -ALIGN=45 -MODE=LOCAL -OUTFMT=pro
-NORM-ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09831458.acn.1.1.1616 -runat_05122002_103421_16305 -NPRO=6 -TPRO 3
-NO_XLPHY -NO_MMAR -THRESEVERY -NEG_SCORES=0 -WAIT -L=0 -I=0 -I=0 -I=0 -I=0
-MARKTIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-VGAPOP=60 -VGAPEXT=60 -DELop=6 -DELtext=7

Database:

GenBml: *
1: qb.ba:*
2: qb.htg:*
3: qb.in:*
4: qb.om:*
5: qb.ov:*
6: qb.pat:*
7: qb.ph:*
8: qb.pl:*
9: qb.pr:*
10: qb.ro:*
11: qb.sts:*
12: qb.sy:*
13: qb.un:*
14: qb.vl:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.ov:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

29: em.vl:*
30: em.htg.hum:*
31: em.htg.in:*
32: em.htg.other:*
33: em.htg.mu:*
34: em.htg.ph:*
35: em.htg.ro:*
36: em.htg.un:*
37: em.htg.vl:*
38: em.sy:*
39: em.htg.hum:*
40: em.htg.mu:*
41: em.htg.other:*

Pred. No. is the number of results predicted by chance 1
score greater than or equal to the score of the result.
and is derived by analysis of the total score distribution

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Accession
1	168	51.7	1643	6	AX287118
2	168	51.7	1643	9	AB015629
3	166	51.1	1200	9	AY042244
4	166	51.1	1800	9	AF290867
5	153	47.1	1510	9	AF245219
6	153	47.1	1535	9	AY042236
7	153	47.1	1548	9	AY042245
8	138	42.5	792	9	AY042240
9	138	42.5	999	9	AY042238
10	125	38.5	1472	9	AY042239
11	125	38.5	4506	9	AF209480S2
12	125	38.5	143619	9	AC008812
13	120	36.9	811	9	AY042237
14	115	35.4	190628	2	AC024709
15	97	29.8	2991	9	AK024750
16	78	24.0	347	9	AY042233
17	78	24.0	939	9	AY042234
18	78	24.0	994	9	AY042232
19	78	24.0	1083	9	AY042226
20	78	24.0	1105	9	AY042231
21	78	24.0	1143	9	AY042225
22	78	24.0	1197	9	AY042222
23	78	24.0	1212	6	AR166109
24	78	24.0	1212	6	AX42004
25	78	24.0	1215	9	AX049404
26	78	24.0	1215	9	AY042221
27	78	24.0	1243	9	AX042230
28	78	24.0	1312	6	AX287104
29	78	24.0	1312	9	HDMLECTIN
30	78	24.0	1315	9	AY042227
31	78	24.0	4266	9	AF290866
32	78	24.0	4424	9	AF209479
33	78	24.0	143619	9	AC008812
34	78	24.0	177962	9	AC008764
35	69	21.2	807	9	AY042227
36	43	13.2	1212	6	AR166106
37	43	13.2	1212	6	AX42004
38	43	13.2	1212	9	AF040865
39	40	12.3	1146	9	AF040755
40	40	12.3	1166	9	AY040319
41	35	10.8	1212	6	AR166107
42	35	10.8	1212	6	AX42004
43	34	10.5	1146	9	AF443727
44	26	8.0	1215	9	AF040866
45	21	6.5	507	9	AY042224

RESULT 1

ALIGNMENTS

AX:87118
 LOCUS AX297118 1643 bp DNA linear PAT 21 NOV 2001
 DEFINITION Sequence 16 from Patent WO0164752.
 ACCESSION AX287118
 VERSION AX287118.1 GI:17049092
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE
 1. Littman, D. P., Kwon, D., van Kester, Y. C., and Goldfarb, R. H. Antibody inhibiting the binding between gp120 and dectin and screening methods. Patent: WO 0164752-A (b. 07-SEP-2001).
 JOURNAL NEW YORK UNIVERSITY (US) ; KATHOLIEKE UNIVERSITEIT NIMHGEN (NL)
 FEATURES
 SOURCE
 1..1643
 Location/Qualifiers
 CDS
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 25..915
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD12298.1"
 /db_xref="GI:17049093"
 /translation="MSDSKEPRVQDILGICGHALVLLSLMILAGVVALLVV SKVPSLSQESQSDAIYONLTOLKAAGVLSKSLQEIYQELTOLKAAGVLPKS KLOEIYQELTRKAAGVLPKSLQEIYQELTRKAAGVLPKSLQEIYQELTRL KAAGVLPKSLQEIYQELTRKAAGVLPKSLQEIYQELTRKAAGVLPKSLQEIYQELTRK KQOITQELTRKAAGVLPKSLQEIYQELTRKAAGVLPKSLQEIYQELTRK VIKTAEQDLPVLEQWRQO"
 BASE COUNT 388 a 429 c 440 g 386 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 4,446-166 Length: 1643
 Score: 168.00 Matches: 168
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 51.69% Indels: 0
 DB: 6 Gaps: 0
 US-09-831-458A-12 (1-325) x AX287118 (1-1643)
 QY 1 Met Ser Asp Ser Lys Glu Pro Arg Val Glu Ileu Glu Leu Leu Glycyl His 20
 DB 25 ATGAGTGACATCCAGAGAACCAAGAGTGTAGTACCTGGGCTCTGCGGCTGAGT 84
 QY 21 GVALALEUVALLEUULEUULEUUSERPHEMETLEUALEAGLYVALLEUVALALE 40
 DB 85 GGGGCGCTGGTGTGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 144
 QY 41 LEUVALGLNVALSERLYSVALPROSERSERLEUSERGGLUGLINSERGLGLNASPALA 60
 DB 145 CTTGTCCCAATGTCTCAAGGCTCCCTCACTCTCTCAAGTCAAGAACCAATCCAGAAC 204
 QY 61 LLETYGLIASHLEUTHTRGLINLEUYSALALEVALGLYGLULEUSERGLYSSERTYS 80
 DB 205 APTATTAATTAATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 264
 QY 81 LGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 100
 DB 265 GGGGCGCTGGTGTGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 144
 QY 101 LYSSELYSLEUGLINSGLILETYGSLINLEUTHTRARGLEUYSALALEVALGLYGLU 120
 DB 325 AATATCAAGATGTAGAGATGTAGAGATGTAGAGATGTAGAGATGTAGAGATGTAGAG 184
 QY 121 LEUPROGLULYSSELYSLEUGLINSGLILETYGSLINLEUTHTRARGLEUYSALALE 140
 DB 385 TTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 444
 QY 141 VALGLYGLULEUPROGLULYSSELYSLEUGLINSGLILETYGSLINLEUTHTRARGLEU 160

DB 445 GGGGCGCTGGTGTGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 504
 QY 161 LYSALALEVALGLYGLULEUPRO 168
 DB 505 AAGGCTGTAGAGATGTAGAGATGTAGAGATGTAGAGATGTAGAGATGTAGAGATGTAG 528
 RESULT 2
 AB015629
 LOCUS
 DEFINITION Homo sapiens mRNA for type II membrane protein similar to HIV gp120-binding C-type lectin, complete cds. clone: HPO1347.
 ACCESSION AB015629
 VERSION AB015629.1 GI:4586835
 KEYWORDS type II membrane protein similar to HIV gp120-binding C-type lectin.
 SOURCE Homo sapiens liver cDNA to mRNA, clone: lib:pkat-liver
 clone: HPO1347.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE
 1. Yokoyama-Kobayashi, M., Yamaguchi, T., Seifried, S., and Kato, S. Selection of cDNAs encoding putative type II membrane proteins on the cell surface from a human full-length cDNA bank
 Gene 228 (1-2): 151-162 (1999)
 JOURNAL MEDLINE
 99173880
 2 (bases 1 to 1643)
 AUTHORS Kato, S.
 TITLES Direct Submission
 JOURNAL Submitted (16-JUN-1998) Seishi Kato, Research Institute of National Rehabilitation Center for the disabled, Department of Rehabilitation Engineering, 4-1 Namiki, Tokorozawa, Saitama 359-8555, Japan (E-mail: seishi.kato@rehab.go.jp, Tel: 042-995-3100 (ex. 2568), Fax: 042-995-3152)
 FEATURES
 SOURCE
 1..1643
 Location/Qualifiers
 CDS
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HPO1347"
 /translation="MSDSKEPRVQDILGICGHALVLLSLMILAGVVALLVV SKVPSLSQESQSDAIYONLTOLKAAGVLSKSLQEIYQELTOLKAAGVLPKS KLOEIYQELTRKAAGVLPKSLQEIYQELTRKAAGVLPKSLQEIYQELTRL KAAGVLPKSLQEIYQELTRKAAGVLPKSLQEIYQELTRKAAGVLPKSLQEIYQELTRK KQOITQELTRKAAGVLPKSLQEIYQELTRKAAGVLPKSLQEIYQELTRK VIKTAEQDLPVLEQWRQO"
 BASE COUNT 388 a 429 c 440 g 386 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 4,446-166 Length: 1643
 Score: 168.00 Matches: 168
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 51.69% Indels: 0
 DB: 6 Gaps: 0
 US-09-831-458A-12 (1-325) x AB015629 (1-1643)
 QY 1 Met Ser Asp Ser Lys Glu Pro Arg Val Glu Ileu Glu Leu Leu Glycyl His 20
 DB 25 ATGAGTGACATCCAGAGAACCAAGAGTGTAGTACCTGGGCTCTGCGGCTGAGT 84
 QY 21 GVALALEUVALLEUULEUULEUUSERPHEMETLEUALEAGLYVALLEUVALALE 40
 DB 85 GGGGCGCTGGTGTGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 144

PEKSLQETIYQETUAKAVGELPQSKQYIYQETITLKIATFEELCKHCKPMTHD
GNYEWSNBNHDSVTATQVPAULVITKAFRNFQIQTSSNPSWMLSDN
QESTWQVDSPLSPSPQRYWNSGPNNSNIFYCAFFSSNNINICIDVNYICKRP
AACFRDE
86.169
/gene="abc"
/note="encodes transmembrane domain"
exon
539.1108
/gene="abc"
/note="encodes coiled-coil neck"
repeat_region
613.1095
/note="encodes likely alpha helical coiled-coil region
that forms the neck of the protein"
/rpt_type=tandem
/rpt_unit=613.681
1557.1708
/gene="abc"
/note="encodes C-type lectin domain CR1 involved in the
formation of the carbohydrate recognition domain"
exon
2417.2529
/gene="abc"
/note="C-type lectin domain CPT2 involved in the
formation of the carbohydrate recognition domain"
exon
3740.13945
/gene="abc"
/note="encodes C-type lectin domain CR3 involved in the
formation of the carbohydrate recognition domain"
exon
4275.24506
/number=7
/number=8
/gene="abc"
exon
BASE COUNT 1106 a 1091 c 1291 g 1017 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 1 090-120 Length: 4506
Score: 125.00 Matches: 125
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 38.46% Indels: 0
DB: 9 Gaps: 0
US-09-831-458a-12 (1-325) x AF209480S2 (1-4506)
QY 44 ValSerLysValProSerSerLeuSerGlnGlnSerGlnGlnAspAlaLeuTyrGln 63
|||||
Db 538 GTGTCCAAAGTCCCGACGTCCCTAAGTCAGAACCAATCCGACAGACGCAATCTACCA 597
QY 64 AsnLeuThrGlnLeuLysAlaAlaValGlyGlnLeuSerGlnLysSerGlnPro 83
|||||
Db 598 AACCTGACCCAGCTTAAAGCTGACAGTGGAGAGTCTCAAGASAAATCTAGCTGAG 657
QY 84 IleTyrGlnLeuLeuThrGlnLeuLysAlaAlaValGlyGlnLeuSerGlnLys 103
|||||
Db 658 ATTACACAGAGGTGACCGACGCTGCAAGCTGCAATGGTGATTTCCCAATCAAAATTC 717
QY 104 LeuGlnGlnLeuTyrGlnGlnLeuThrArgLeuLysAlaValGlyGlnLeuSer 123
|||||
Db 718 GTGACGAGATCTACGACGATCTACGACGATCTACGACGATCTACGACGATCTACG 777
QY 124 LysSerLysLeuGlnGlnLeuTyrGlnGlnLeuThrArgLeuLysAlaValGlyGln 143
|||||
Db 778 AAATCCAAAGCTGCGAGATCTACGACGATCTACGACGATCTACGACGATCTACG 837
QY 144 LeuProGlnLysSerLysLeuGlnGlnLeuTyrGlnGlnLeuThrArgLeuLysAla 163
|||||
Db 838 TTGGCAGAGAAATCCAAAGCTGCGAGATCTACGACGATCTACGACGATCTACG 897
QY 164 ValGlyGlnLeuPro 168

|||||
Db R98 GTGGTCAATTCGA 912
RESULT 12
AC008812
LOCUS AC008812 144619 bp DNA linear
DEFINITION Homo sapiens chromosome 19 clone CH2102P19, cont.
AC008812
ACCESSION AC008812
VERSION AC008812.7 GI:9558577
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata
Mammalia; Eutheria; Primates; Catarrhini; Hominidae
1 (bases 1 to 143619)
DOE Joint Genome Institute and Stanford Human Ge
Title Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 143619)
DOE Joint Genome Institute
AUTHORS Direct Submission
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 143619)
Genome Institute, 2800 Mitchell Drive, Walnut Cr
3 (bases 1 to 143619)
DOE Joint Genome Institute and Stanford Human Ge
Title Direct Submission
JOURNAL Unpublished
REFERENCE 4 (bases 1 to 143619)
DOE Joint Genome Institute
AUTHORS Direct Submission
TITLE Direct Submission
JOURNAL Unpublished
COMMENT Submitted (04 AUG 1997) Production Sequencing for
Genome Institute, 2800 Mitchell Drive, Walnut Cr
On Jul 28, 2000 this sequence version replaced the
Draft Sequence Produced by DOE Joint Genome Inst
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Ce
www.sngc.stanford.edu
Quality: Phrap quality 99.99% of sequence
Estimated Total Number of Errors is 3.8.
Location/Qualifiers
1. 143619
/organism "Homo sapiens"
/db_xref "taxon:9606"
/chromosome "19"
/clone "CH2102P19"
/feature "CDS" (1-143619)
BASE COUNT 9628 a 35385 c 34114 g 34192 t
ORIGIN
Alignment Scores:
Pred. No.: 1 325-119 Length: 14361
Score: 125.00 Matches: 125
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 38.46% Indels: 0
DB: 9 Gaps: 0
US-09-831-458a-12 (1-325) x AC008812 (1-143619)
QY 44 ValSerLysValProSerSerLeuSerGlnGlnSerGlnGlnAspAlaLeuTyrGln 63
|||||
Db 68748 GTGTCCAAAGTCCCGACGTCCCTAAGTCAGAACCAATCCGACAGACGCAATCTACCA 657
QY 64 AsnLeuThrGlnLeuLysAlaAlaValGlyGlnLeuSerGlnLysSerGlnPro 83
|||||
Db 68808 AACCTGACCCAGCTTAAAGCTGACAGTGGAGAGTCTCAAGASAAATCTAGCTGAG 657
QY 84 IleTyrGlnLeuLeuThrGlnLeuLysAlaAlaValGlyGlnLeuSerGlnLys 103
|||||
Db 68868 ATTACACAGAGGTGACCGACGCTGCAAGCTGCAATGGTGATTTCCCAATCAAAATTC 717
QY 104 LeuGlnGlnLeuTyrGlnGlnLeuThrArgLeuLysAlaValGlyGlnLeuSer 123
|||||
Db 68928 GTGACGAGATCTACGACGATCTACGACGATCTACGACGATCTACGACGATCTACG 777
QY 124 LysSerLysLeuGlnGlnLeuTyrGlnGlnLeuThrArgLeuLysAlaValGlyGln 143
|||||
Db 68988 AAATCCAAAGCTGCGAGATCTACGACGATCTACGACGATCTACGACGATCTACG 837
QY 144 LeuProGlnLysSerLysLeuGlnGlnLeuTyrGlnGlnLeuThrArgLeuLysAla 163
|||||
Db 68988 AAATCCAAAGCTGCGAGATCTACGACGATCTACGACGATCTACGACGATCTACG 897

Q7	144	Leptodermyssestysseuincinguletyringinguleutrravglouysaliaia	163
D9	69048	TTGSCAAGAAATCTACTTTCAGACGATTACCTAGGAATCGACCTGGGTGGAACTCTGCA	69107
Q7	164	Valsgylcinleupro	168
D9	69108	GCGCATGATTTCCCA	69122
<hr/>			
RESULT 13			
A0042237			
DEFINITION	AY042237	Homo sapiens MDG-SIGN2 type VI isoform (ClD209L) mRNA, complete cds.	811 bp mRNA linear pri 30-AUG-2001
ACCESSION	AY042237		
VERSION	AY042237		
KEYWORDS	AY042237.1	GI:15383611	
SOURCE			
ORGANISM			Homo sapiens.
REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS			Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE			1 (bases 1 to 811)
JOURNAL			Mummidi,S., Cattano,G., Lam,I., Boefle,A., Telles,V., Begum,K.,
MEDLINE			Jimenez,F., Ahuja,S.S. and Ahuja,S.K
PUBMED			extensive repertoire of membrane-bound and soluble dendritic
REFERENCES			cell-specific TBM-3-grabbing nonintegrin 1 (DG-SIGN) and DG-SIGN
AUTHOPS			isoforms, inter individual variation in expression of DG-SIGN
TITLE			transcripts
JOURNAL			J. Biol Chem. 276 (43), 33196-33212 (2001)
MEDLINE			21413847
PUBMED			11337487
REFERENCES			2 (bases 1 to 811)
AUTHOPS			Mummidi,S. and Ahuja,S.K.
TITLE			Submitted (03-JUN-2001) Medicine, South Texas Veterans Health Care
JOURNAL			System and University of Texas Health Science Center at San
FEATURES			Antonio, 7703 Floyd Curl Drive, San Antonio, TX 78229, USA
source			Location/Qualifiers
			1..811
			/organism="Homo sapiens"
			/db_xref="taxon:9606"
			/chromosome="19"
			/map_ "19p13.3"
			1..811
			/gene="CD209L"
			1..699
			/note="similar to the nucleotide sequence deposited in
			GenBank Accession Number AC009812; alternatively spliced"
			/codon_start=1
			/product="MDG-SIGN2 type VI isoform"
			/protein_id="AAK91862.1"
			/db_xref="GI:15383612"
			/translation="MSDSKPPVQGLLEDDPTTSGRIFPRDFEQQLIHKKSSST
			CGHSAIVDLTSLFMLAGVALIIVGVSPSSLQEOSEQUALIYONILYLKAAGVG
			LSEKSALVEITVELTKAVNGELPEKSKLOEITVELTRIKAAVFETPKRSIDETLY
			FITLKAAVPRTCPHCPRKWTFTEFGNCTFPNSNCQPMHNSTVAQCVVAQIVITTA
			EQLPATEQWKQTG*

[illegible]

```

* be preserved.
1 1244: contig of 1244 bp in length
1245 1344: gap of unknown length
1345 2396: contig of 1052 bp in length
2397 2497: gap of unknown length
2497 3686: contig of 1089 bp in length
3686 3686: gap of unknown length
3686 5159: contig of 1473 bp in length
5159 5259: gap of unknown length
5259 6724: contig of 1366 bp in length
6724 6725: gap of unknown length
6725 8402: contig of 1677 bp in length
8402 8502: gap of unknown length
8502 10173: contig of 1671 bp in length
10173 10273: gap of unknown length
10273 12510: contig of 2238 bp in length
12511 12610: gap of unknown length
12611 14735: contig of 2125 bp in length
14736 14836: gap of unknown length
14836 16443: contig of 1608 bp in length
16443 16543: gap of unknown length
16543 19269: contig of 2726 bp in length
19270 19370: gap of unknown length
19370 21791: contig of 2321 bp in length
21791 25544: contig of 3753 bp in length
25544 25644: gap of unknown length
25644 28329: contig of 2585 bp in length
28329 31740: gap of unknown length
31740 31740: contig of 3311 bp in length
31740 36466: gap of unknown length
36466 36567: contig of 4727 bp in length
36567 36567: gap of unknown length
36567 39593: contig of 3027 bp in length
39593 39594: gap of unknown length
39594 39594: gap of unknown length
39594 46098: contig of 6405 bp in length
46098 46199: gap of unknown length
46199 50495: contig of 4296 bp in length
50495 50495: gap of unknown length
50495 56367: contig of 5773 bp in length
56367 56367: gap of unknown length
56367 61903: contig of 5436 bp in length
61903 62004: gap of unknown length
62004 62004: contig of 5156 bp in length
62004 67260: gap of unknown length
67260 74171: contig of 6912 bp in length
74171 74172: gap of unknown length
74172 80579: contig of 6308 bp in length
80579 80580: gap of unknown length
80580 88540: contig of 7861 bp in length
88540 88541: gap of unknown length
88541 95519: contig of 6878 bp in length
95519 95519: gap of unknown length
95519 102484: contig of 6866 bp in length
102484 102485: gap of unknown length
102485 102585: contig of 6884 bp in length
102585 109469: gap of unknown length
109469 130222: contig of 20554 bp in length
130223 130323: gap of unknown length
130323 130323: contig of 24567 bp in length
130323 154890: gap of unknown length
154890 171709: contig of 16720 bp in length
171710 171810: gap of unknown length
171810 171810: contig of 18819 bp in length.

FEATURES
source
1. 190628
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/clone="RP11-42J18"
1. 1244
/note="assembly_name:Contig15"
1345. 2396
/note="assembly_name:Contig15"

misc_feature
2497..3585
/note="assembly_name:Contig17"
3686..5158
/note="assembly_name:Contig20"
5259..6624
/note="assembly_name:Contig23"
6725..8401
/note="assembly_name:Contig24"
8502..10172
/note="assembly_name:Contig25"
10273..12510
/note="assembly_name:Contig29"
12511..14735
/note="assembly_name:Contig30"
14836..16443
/note="assembly_name:Contig31"
16544..19269
/note="assembly_name:Contig32"
19370..21790
/note="assembly_name:Contig33"
21791..25543
/note="assembly_name:Contig34"
25644..28328
/note="assembly_name:Contig35"
28329..31739
/note="assembly_name:Contig36"
31740..36465
/note="assembly_name:Contig37"
36567..39592
/note="assembly_name:Contig38"
39594..46097
/note="assembly_name:Contig39"
46199..50494
/note="assembly_name:Contig40"
50595..56366
/note="assembly_name:Contig41"
56368..61902
/note="assembly_name:Contig42"
62004..67159
/note="assembly_name:Contig43"
67260..74170
/note="assembly_name:Contig44"
74272..80578
/note="assembly_name:Contig45"
80680..88539
/note="assembly_name:Contig46"
88641..95518
/note="assembly_name:Contig47"
95619..102483
/note="assembly_name:Contig48"
102585..109468
/note="assembly_name:Contig49"
109569..130221
/note="assembly_name:Contig50"
130324..154889
/note="assembly_name:Contig51"
154990..171708
/note="assembly_name:Contig52"
171811..190627
/note="assembly_name:Contig53"

BASE COUNT 43336 a 49517 c 49715 g 44933 t 3127 other
ORIGIN
Alignment Scores:
Pred. No.: 4,93e-109 Length: 190628
Score: 115.00 Matches: 115
Percent Similarity: 100.00% Conserved: 0
Best Local Similarity: 100.00% Mismatch: 0

```


JOURNAL

Submitted (18-JUN-2001) Department of Medicine, South Texas
Veterans Health Care System and University of Texas Health Science
Center, 7703, Floyd Curl Drive, San Antonio, TX 78296, USA

FEATURES

Source

1..547
/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="19p13.3"

1..547
/gene="CD209"

1..126
/gene="CD209"

/note="Putative open reading frame; encodes truncated
protein tpe-SIGN1B"

/evidence="not-experimental"

101..547

/gene="CD209"

/note="Similar to the nucleotide sequence deposited in
Genbank Accession Number AC008812; alternatively spliced"

/codon_start=1

/product="SDC-SIGN1B type IV isoform"

/protein_id="AAK91858.1"

/db_xref="GI:15281097"

/translation="MASACGSDPFTSHSEPOLRRIHQPIRGYSIAVSKVPSIS
QEOSRDATYONLTOLKAAGVGLSEKSLQEIYQELTOLKAAGVGLPEKSKIQEITOE
LTRKAAGVGLPEKSKIQEITOEITRKAAGVGLPEKSKIQEITOE"

BASE COUNT

144 a 136 c 167 g 100 t

Alignment Scores:

Prod. No.: 4 38a-72 Length: 547
Score: 78.00 Matches: 78
Percent Similarity: 100.00% Conserved: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 24.00% Indels: 0
DB: 9 Gaps: 0

US-09-831-458a-12 (1-325) x AY042233 (1-547)

QY 58 GINASPALIAIETRYGLASNTLEUTHRTGIDLEULYSALIAIAVALIGLULSERGLU 77

DB 248 CAAGACGCGACTACCAAGACCTGACCAAGCTTAAGGCTGACGCTGACCTGACG 307

QY 78 TYSSELYSLGGLGGLIETRYGLINGLIETRYGLINGLIETRYGLINGLIETRYGLIN 97

DB 308 AATCCAGAGCTGACGAGATATACAGAGAGCTGACAGAGAGCTGAGAGAGCTGAG 367

QY 98 LEUPROCIULYSSELYSLGGLGGLIETRYGLINGLIETRYGLINGLIETRYGLIN 117

DB 368 CTTCAGAGAAATCTAAGCTGACAGATATACAGAGAGAGCTGAGAGAGCTGAG 427

QY 118 VALIGLULLEUPROCIULYSSELYSLGGLGGLIETRYGLINGLIETRYGLIN 135

DB 428 GTGGCTGAGCTTCCAGAAATCTAAGCTGAGAGAGCTGAGAGAGCTGAG 481

RESULT 17

LOCUS AY042233 939 bp mRNA linear FRI 28-AUG-2001

DEFINITION Homo sapiens SDC-SIGN1A type III isoform (CD209) mRNA, complete

VERSION AY042233

KEYWORDS AY042233.1 GI:15281076

SOURCE Homo sapiens.

ORGANISM Homo sapiens

REFERENCE Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:

1 (bases 1 to 939)

Mammalia: Eutheria: Primates: Catarrhini, Hominoidea: Hominidae: Homo

Mummid, S., Catano, G., Lam, L., Hoeft, A., Telles, V., Begum, K.,

extensive repertoire of membrane-bound and soluble dendritic

cell-specific ICAM-3-grabbing nonintegrin 1 (DC-SIGN) and DC-SIGN2

isoforms. Inter-individual variation in expression

transcripts

JOURNAL 1. Biol. Chem. 276 (25): 2199-2212 (2001)

REFERENCE 2 (bases 1 to 939)

Mummid, S., and Abuja, S.K.

Direct Submission

Submitted (18-JUN-2001) Department of Medicine,

Veterans Health Care System and University of

Center, 7703, Floyd Curl Drive, San Antonio, TX

Location/Qualifiers

1..939

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="19"

/map="19p13.4"

1..939

/gene="CD209"

1..939

/note="Similar to the nucleotide sequence

Genbank Accession Number AC008812; alternatively spliced"

/codon_start=1

/product="SDC-SIGN1A type III isoform"

/protein_id="AAK91848.1"

/db_xref="GI:15281077"

/translation="MASACGSDPFTSHSEPOLRRIHQPIRGYSIAVSKVPSIS
QEOSRDATYONLTOLKAAGVGLSEKSLQEIYQELTOLKAAGVGLPEKSKIQEITOE
LTRKAAGVGLPEKSKIQEITOEITRKAAGVGLPEKSKIQEITOE"

BASE COUNT 239 a 260 c 257 g 183 t

Alignment Scores:

Prod. No.: 6 47a-72 Length: 939
Score: 78.00 Matches: 78
Percent Similarity: 100.00% Conserved: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 24.00% Indels: 0
DB: 9 Gaps: 0

US-09-831-458a-12 (1-325) x AY042233 (1-939)

QY 58 GINASPALIAIETRYGLASNTLEUTHRTGIDLEULYSALIAIAVALIGLULSERGLU 77

DB 220 CAAGACGCGACTACCAAGACCTGACCAAGCTTAAGGCTGACGCTGACG 307

QY 78 TYSSELYSLGGLGGLIETRYGLINGLIETRYGLINGLIETRYGLINGLIETRYGLIN 97

DB 480 AATCCAGAGCTGACGAGATATACAGAGAGCTGACAGAGAGCTGAGAGAGCTGAG 367

QY 98 LEUPROCIULYSSELYSLGGLGGLIETRYGLINGLIETRYGLINGLIETRYGLIN 117

DB 340 CTTCAGAGAAATCTAAGCTGACAGATATACAGAGAGAGCTGAGAGAGCTGAG 427

QY 118 VALIGLULLEUPROCIULYSSELYSLGGLGGLIETRYGLINGLIETRYGLIN 135

DB 400 GTGGCTGAGCTTCCAGAAATCTAAGCTGAGAGAGCTGAGAGAGCTGAG 481

RESULT 18

LOCUS AY042232 934 bp mRNA linear FRI 28-AUG-2001

DEFINITION Homo sapiens SDC-SIGN1B type III isoform (CD209)

VERSION AY042232

KEYWORDS AY042232.1 GI:15281094

SOURCE Homo sapiens.

ORGANISM Homo sapiens

REFERENCE Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:

1 (bases 1 to 939)

Mammalia: Eutheria: Primates: Catarrhini, Hominoidea: Hominidae: Homo


```

DB      208  CTTCTTACGAATCTAATCTGACGAGAT-TATTAATTAATCTAAATGAGATGAGAA 207
QY      118  ValGlyGluLeuProGlyLysSerLysLeuGlnGlnIleTyrGlnGlnLeuThr 135
Db      268  GTGGGTACCTTCCACAGCAAACTTAAGCTGTGAGATATCTATGAGAGAGGTTGAGT 327

RESULT 20
AY042231      1105 bp      mRNA      linear      PE1 24-AUG-2001
LOCUS        Homo sapiens CDC-SIGNIB type II isoform (CD209) mRNA, complete cds,
DEFINITION   alternatively spliced.
ACCESSION   AY042231
VERSION      AY042231.1 GI:15281092
KEYWORDS
SOURCE       Homo sapiens.
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 1105)
AUTHORS      Mummidi,S., Catano,G., Lam,L., Hoelle,A., Telles,V., Sejum,K.,
             Jimenez,F., Ahuja,S.S. and Ahuja,S.K.
             Extensive repertoire of membrane-bound and soluble dendritic
             cell-specific ICAM-3-grabbing nonintegrin 1 (DC-SIGN) and DC-SIGN2
             isoforms: inter-individual variation in expression of DC-SIGN
             transcripts
JOURNAL      J. Biol. Chem. 275 (45), 33196-33212 (2001)
MEDLINE      21413647
PUBMED       11337487
REFERENCE    2 (bases 1 to 1105)
AUTHORS      Mummidi,S. and Ahuja,S.K.
TITLE        Direct Submission
JOURNAL      Submitted (18-JUN-2001) Department of Medicine, South Texas
             Veterans Health Care System and University of Texas Health Science
             Center, 7703, Floyd Curl Drive, San Antonio, TX 78229, USA
FEATURES
SOURCE       1..1105
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /chromosome="19"
             /map="19p13.3"
             /gene="CD209"
             /gene="CD209"
             /note="putative open reading frame; encodes truncated
             protein tDC-SIGNIB"
             /evidence="not_experimental"
             /codon_start=1
             /product="SDC-SIGNIB type II isoform"
             /protein_id="AAK91856.1"
             /db_xref="GI:15281093"
             /translation="MASACPGSDFTSIHSEFQRLQIFPRTGCGSLAVSKVPSIS
             QEGSRDAIYONITOLKAAGVGLSEKSKLOEIVQELTOLKAAGVGLPEKSKOELIYOE
             LTRKAAGVGLPEKSKLOEIVQELTOLKAAGVGLPEKSKMGELIYELTRKAAGVGLP
             EKSKOELIYOEIVQELTOLKAAGVGLPEKSKMGELIYOEIVQELTOLKAAGVGLPEK
             AQLVTKSAKQVNFLOLQSSKSNFTLMTLSLNLQFGLTQWVWTSKLLPSKQVWNPQ
             EPNNVGEEDCAEPSCGNMDKCNLAKFWIKKSAASCSDPEDFLSPAPATPNPPA"
BASE COUNT  288 a      289 c      321 g      207 t
ORIGIN
Alignment Scores:
Pred. No.:      7,28e-72      Length:      1105
Score:          78.00         Matches:      78
Percent Similarity: 100.00%   Conservat.:  0
Best Local Similarity: 100.00% Mismatches:  0
Query Match:    24.00%       Indels:      0
DB:             9            Gaps:         0

```

```

US-09-831-458A 12 (1 325) x AY042231 (1-1105)
CY      58  GlnSPAlaIleTyrGlnAsnLeuThrGlnGlnGlnGlnAlaValGlyGln
Db      248  CACAGCTGATCTATCTAAATCTAAATCTAAATCTAAATCTAAATCTAAATCTAAATCTAA 327
QY      78  LysSerLysLeuGlnGlnIleTyrGlnGlnLeuThrGlnGlnGlnAlaVal
Db      308  AAATCTAAATCTGACGAGATCTAATTAATTAATCTAAATCTGACGAGATCTGAG 327
QY      98  LeuProGlyLysSerLysLeuGlnGlnIleTyrGlnGlnLeuThrArgLeu
Db      368  CTTCTTACGAATCTAATCTGACGAGAT-TATTAATTAATCTAAATGAGATGAGAA 327
QY      118 ValGlyGluLeuProGlyLysSerLysLeuGlnGlnIleTyrGlnGlnLeuThr 135
Db      428  GTGGGTACCTTCCACAGCAAACTTAAGCTGTGAGATATCTATGAGAGAGGTTGAGT 327

RESULT 21
AY042225      1143 bp      mRNA      linear      PE1 20-1
LOCUS        Homo sapiens SDC-SIGNIA type I isoform (CD209) mR,
DEFINITION   alternatively spliced.
ACCESSION   AY042225
VERSION      AY042225.1 GI:15281080
KEYWORDS
SOURCE       Homo sapiens.
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
Mammalia; Eutheria; Primates; Catarrhini; Hominid.
REFERENCE    1 (bases 1 to 1143)
AUTHORS      Mummidi,S., Catano,G., Lam,L., Hoelle,A., Telles,V.,
             Jimenez,F., Ahuja,S.S. and Ahuja,S.K.
             Extensive repertoire of membrane-bound and soluble
             cell-specific ICAM-3-grabbing nonintegrin 1 (DC-SIGN)
             isoforms: inter-individual variation in expression
             transcripts
JOURNAL      J. Biol. Chem. 275 (45), 33196-33212 (2001)
MEDLINE      21413647
PUBMED       11337487
REFERENCE    2 (bases 1 to 1143)
AUTHORS      Mummidi,S. and Ahuja,S.K.
TITLE        Direct Submission
JOURNAL      Submitted (18-JUN-2001) Department of Medicine, S
             Veterans Health Care System and University of Tex
             Center, 7703, Floyd Curl Drive, San Antonio, TX 7
             Location/Qualifiers
FEATURES
SOURCE       1..1143
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /chromosome="19"
             /map="19p13.3"
             /gene="CD209"
             /gene="CD209"
             /note="similar to the nucleotide sequence
             GenBank Accession Number AC008812; alter
             /codon_start=1
             /product="SDC-SIGNIA type I isoform"
             /protein_id="AAK91850.1"
             /db_xref="GI:15281081"
             /translation="MSISKPEKRLQGLLEFLQLQGLF
             QEGSRDAIYONITOLKAAGVGLSEKSKLOEIVQELTOLKA
             LTRKAAGVGLPEKSKLOEIVQELTOLKAAGVGLPEKSKM
             EKSKOQEIYOEIVQELTOLKAAGVGLPEKSKOQEIYOEIVQ
             TOLKAAGVGLPEKSKOQEIYOEIVQELTOLKAAGVGLPEK
             NPLQLOSRSNRFIMLSDINQEGTQWVWVTSKLLPSKQVWNP
             FSGNGNMDKCTNIAFWIKKSAASCSPREDFLSIAPATPNPPA"
BASE COUNT  309 a      296 c      466 g      208 t
ORIGIN

```


gene
misc_feature
CDS

/map="1p13.3"
1..1315
/gene="CD209"
1..126
/note="putative open reading frame; encodes truncated protein tdc-sicn1b"
/evidence="not_experimental"
101..1315
/gene="CD209"
/note="Similar to the nucleotide sequence deposited in GenBank Accession Number AC086122, alternative splicing of product_start=1
/product="MPC SIGN1B type I isoform"
/protein_id="AAK91854.1"
/db_xref="GI:15281080"
/translation="MASSAPOSTSTSHSEFQPSLHPGPKYKSLAGTGHPIV
LQLSPFLAGLVOVSKVPSISOBOSKODATYONLTOLKAAGELSEKSLQELVQ
ELTOLKAAGVLPKSKLOETLQETLRIKAAGVLPKSKLOETLQETLRIKAAGVLP
PEKSKMOETLQETLRIKAAGVLPKSKLOETLQETLRIKAAGVLPKSKLOETLQ
LRIKAAGVLPKSKLOETLQETLQKAVERLCHPCWMTFPGNCEYFMSNSGN
WHDSTTAKKQVADIVIKSAEONFLQVSSKSNKPTWMLSLNBDTHOWVDSPL
LPSRKQYWNKQEPNNVTEFCARFSGNMNMKCNLAKFMTIKKSAASSPDEQFL
SPAPATPNPPPA"

BASE COUNT 332 a 345 c 346 j 252 t

ORIGIN

Alignment Scores:
pred. No.: 8.25e-72 Length: 1315
Score: 78.00 Matches: 78
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 24.00% Indels: 0
Dbs: 9 Gaps: 0

US-09-831-458a-12 (1-325) x AF294229 (1-1315)

QY 58 GlnASPAlaIleTyrGlnAsnLeuThrGlnLeuLysAlaIleValGlyLeuSerGln 77
|||||
ID 320 CAAAGACGCGATCTACCAAGACCGAGACCGAGACCGAGACCGAGACCGAGACCGAG 379
QY 78 LysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 97
|||||
Db 380 AAATCCAAATCTAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 439
QY 98 LeuProGlnLysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 117
|||||
Db 440 GTTCCAGAGAAATCTAAGCTCTAGAGATCTAGAGATCTAGAGATCTAGAGATCTAG 499
QY 118 ValGlyGlnLeuProGlnLysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGln 135
|||||
Db 500 GTTCCAGAGAAATCTAAGCTCTAGAGATCTAGAGATCTAGAGATCTAGAGATCTAG 553

RESULT 31
AF290886 4266 bp mRNA linear PRI 30 MAR 2001
DEFINITION Homo sapiens DC-SIGN mRNA, complete cds.
ACCESSION AF290886
VERSION AF290886.1 GI:1332467
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 4266)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Hashitrova, A.A., Geijtenbeek, T.B.H., van Duynhoven, G.C.P., van
Vlier, S.J., Eilering, J.B.G., Martin, M.P., Wu, L., Martin, T.P.,
Vlieh, N., Knolle, P.A., Kwaalman, V.N., van Kooyk, Y. and
Carroll, M.
A dendritic cell-specific intercellular adhesion molecule-
X-tracting nonintegrin (DC-SIGN)-related protein is highly
expressed on human liver sinusoidal endothelial cells and promotes
HIV-1 infection

JOURNAL J. Exp. Med. 193 (9), 671-676 (2001)
MEDLINE 21157496
PUBMED 11257134
REFERENCE 2 (bases 1 to 4266)
AUTHORS Hashitrova, A.A.
TITLE Direct Submersion
JOURNAL Submitted (27-JUN-2000) EMBL, NCBI-EBF, P.O. Box B, Frederick, MD
21702, USA

FEATURES
source
1..4266
/organism="Homo sapiens"
/db_xref="GI:1332467"
/chromosome="19"
/map="1p13.3-p13.2"
101..1224
/cds_start=1
/product="DC-SIGN"
/protein_id="AAK20957.1"
/db_xref="GI:13383468"
/translation="MASSAPOSTSTSHSEFQPSLHPGPKYKSLAGTGHPIV
LQLSPFLAGLVOVSKVPSISOBOSKODATYONLTOLKAAGVLPKSKLOETLQ
ELTOLKAAGVLPKSKLOETLQETLRIKAAGVLPKSKLOETLQETLRIKAAGVLP
PEKSKMOETLQETLRIKAAGVLPKSKLOETLQETLRIKAAGVLPKSKLOETLQ
LRIKAAGVLPKSKLOETLQETLQKAVERLCHPCWMTFPGNCEYFMSNSGN
WHDSTTAKKQVADIVIKSAEONFLQVSSKSNKPTWMLSLNBDTHOWVDSPL
LPSRKQYWNKQEPNNVTEFCARFSGNMNMKCNLAKFMTIKKSAASSPDEQFL
SPAPATPNPPPA"

BASE COUNT 906 a 1170 c 995 g 1195 t

ORIGIN

Alignment Scores:
pred. No.: 1.93e-71 Length: 4266
Score: 78.00 Matches: 78
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 24.00% Indels: 0
Dbs: 9 Gaps: 0

US-09-831-458a-12 (1-325) x AF290886 (1-4266)

QY 58 GlnASPAlaIleTyrGlnAsnLeuThrGlnLeuLysAlaIleValGlyLeuSerGln 77
|||||
Db 239 CAAAGACGCGATCTACCAAGACCGAGACCGAGACCGAGACCGAGACCGAGACCGAG 288
QY 78 LysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 97
|||||
Db 289 AAATCCAAATCTAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAGAGATCTAGAG 348
QY 98 LeuProGlnLysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 117
|||||
Db 349 GTTCCAGAGAAATCTAAGCTCTAGAGATCTAGAGATCTAGAGATCTAGAGATCTAG 408
QY 118 ValGlyGlnLeuProGlnLysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGln 135
|||||
Db 409 GTTCCAGAGAAATCTAAGCTCTAGAGATCTAGAGATCTAGAGATCTAGAGATCTAG 462

RESULT 32
AF294479 4424 bp DNA linear PRI 17 SEP 2000
DEFINITION Homo sapiens probable mannose-binding C-type lectin DC-SIGN gene,
complete cds.
ACCESSION AF294479
VERSION AF294479.1 GI:10179609
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 4424)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Sollieux, E.J., Bartunko, and Townsend, J.
TITLE DC-SIGN: A related gene, DC-SIGN2, and CD23 form a cluster on 1p13
JOURNAL Immunol. 165 (6), 2037-2042 (2000)
MEDLINE 20432267

PUBMED 10975799
 REFERENCE 2 (bases 1 to 4424)
 AUTHORS Solleux E.J.
 TITLE Direct Submission
 JOURNAL Submitted (26-NOV-1999) Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge, Cambs CB2 1QP UK
 FEATURES
 SOURCE
 1..4424
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="19"
 /map="19p13; adjacent to CD25"
 join(42..47,188..247,874..945,1312..1887,2312..2463,3185..3297,4167..4368)
 /product="probable mannose-binding C-type lectin DC-STON"
 join(42..47,188..247,874..945,1312..1887,2312..2463,3185..3297,4167..4368)
 /note="type II integral membrane protein"
 /codon_start=1
 /product="probable mannose-binding C-type lectin DC-STON"
 /protein_id="AA013814.1"
 /db_xref="gi:10179610"
 /translation="MSDSKEPLQQGLLEPEQLGRIPEPTGYSKSLAGLGHPLV
 LLIISFTLLAGLIVQSVSSISQSGKDAIVGNITOLKAAGVELSEKSKDQITQ
 ELITLKAAGVELPEKSKDQITQELITLKAAGVELPEKSKDQITQELITLKAAGVEL
 PEKSKDQITQELITLKAAGVELPEKSKDQITQELITLKAAGVELPEKSKDQITQEL
 LTRIKAAVELEPEKSKDQITQELITLKAAGVELPEKSKDQITQELITLKAAGVELPEKSKDQITQEL
 WHDSITACEVGAQIVLRSAEONFLQSSRSRLPDMGISTDINQGTWQWDGSP
 LIPSPKQVNRGEPNNVGEDECAFEPSGNGMNDKCNLAKFWIKKSAASCSRDEQFL
 SPAPATNPPEPA"
 <42..87
 /note="encodes cytoplasmic domain"
 /number=1
 188..247
 /note="encodes cytoplasmic domain"
 /number=2
 874..945
 /note="encodes transmembrane region"
 /number=3
 1318..1887
 /number=4
 1392..1874
 /note="encodes likely alpha helical coiled-coil region
 that forms the neck of the protein"
 /rpt_type=tandem
 /rpt_unit=1392..1460
 2312..2463
 /note="encodes C-type lectin domain CRD1 involved in the
 formation of the carbohydrate recognition domain"
 /number=5
 3185..3297
 /note="encodes C-type lectin domain CRD2 involved in the
 formation of the carbohydrate recognition domain"
 /number=6
 4167..4368
 /note="encodes C-type lectin domain CRD3 involved in the
 formation of the carbohydrate recognition domain"
 /number=7
 BASE COUNT 1029 a 1027 c 1417 g 951 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,986-71 Length: 4424
 Score: 78.00 Matches: 78
 Percent Similarity: 100.00% Conserved: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 24.00% Indels: 0
 DB: Gaps: 0
 US-09-831-458a-12 (1-325) x AF209479 (1-4424)
 QY 58 GlnAspAlaIleTyrGlnAsnLeuThrGlnIleuLysAlaAlaValGlyCysLeuSArgIu 77
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 1359 CAAGACGATCTACCAAACTGACCTTAAAGCTGACG3812A 418
 QY 78 LysSerLysLeuGlnGlnIleTyrGlnIleuThrGlnIleuLysAlaAla 418
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1419 AAATCCAACTGCAAGAAATCAACAACATCAACCAACATCAAG3812A 418
 QY 98 LeuProGlnLysSerLysLeuGlnGlnIleTyrGlnIleuThrGlnIleuLysAlaAla 418
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1479 CTTCCAGCAAGATCTAAAGCTGACGACATCTACAGACAGCTGAC3812A 418
 QY 118 ValGlyCysLeuProGlnLysSerLysLeuGlnGlnIleTyrGlnIleuThrGlnIleuLysAlaAla 418
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1539 GTGGCTAGCTTCCAGCAAGATCTAAAGCTGACGACATCTACAGACAG3812A 418
 RESULT 33
 AC008812/c 14619 bp DNA 11000
 LOCUS
 DEFINITION Homo sapiens chromosome 19 clone CTD-210ZP19, ccd
 AC008812
 VERSION AC008812.7 GI:9558577
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea
 1 (bases 1 to 143619)
 REFERENCE
 AUTHORS DOE Joint Genome Institute and Stanford Human Ge
 TITLE Direct Submission
 JOURNAL Unpublished
 2 (bases 1 to 143619)
 AUTHORS
 TITLE DOE Joint Genome Institute
 JOURNAL Direct Submission
 Submitted (03-AUG-1999) Production Sequencing Pa
 Genome Institute, 2800 Mitchell Drive, Walnut Cr
 3 (bases 1 to 143619)
 REFERENCE
 AUTHORS DOE Joint Genome Institute and Stanford Human Ge
 TITLE Direct Submission
 JOURNAL Submitted (28-JUL-2000) DOE Joint Genome Insti
 Submitted (28-JUL-2000) DOE Joint Genome Insti
 Drive, Walnut Creek, CA 94598, USA
 On Jul 24, 2000 this sequence version replaced a
 Draft Sequence Produced by DOE Joint Genome Inst
 www.jgi.doe.gov
 Finishing completed at Stanford Human Genome Cen
 www.sshc.stanford.edu
 Quality: Phrap Quality = 40.99, 98.61 Sequence
 Estimated total Number of Errors is 3.8.
 FEATURES
 SOURCE
 1..143619
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="19"
 /clone="CTD-210ZP19"
 BASE COUNT 36628 a 35385 c 45414 g 36192 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,446-79 Length: 14361
 Score: 78.00 Matches: 78
 Percent Similarity: 100.00% Conserved: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 24.00% Indels: 0
 DB: Gaps: 0
 US-09-831-458a-12 (1-325) x AC008812 (1-143619)
 QY 58 GlnAspAlaIleTyrGlnAsnLeuThrGlnIleuLysAlaAlaValGlyCysLeuSArgIu 77
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 49157 CAAGCAAGCTGCAAGAAATCAACAACATCAACCAACATCAAG3812A 418
 QY 78 LysSerLysLeuGlnGlnIleTyrGlnIleuThrGlnIleuLysAlaAla 418
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 49097 AAATCCAACTGCAAGAAATCAACAACATCAACCAACATCAAG3812A 418

FKQYNNRGEPPNNVGEEDCAEFESGNGWMDKCNLAKFWICKKSAASCSRDDEOFLSPAP
ATNPPPA*
BASE COUNT 125 a 142 c 136 g 104 t
ORIGIN

Alignment Scores:

Prod. No.: 2.32e-12 Length: 507
Score: 21.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.46% Indels: 0
DB: 9 Gaps: 0

US-09-831-458A-12 (1-325) x AY042224 (1-507)

UY 253 TRPMTGILYLouSeAspIeUASoGInGluGlyTheIrpCInTrpValAspClySeIPro 272
|||||
DB 245 TGGATGGGACTTTCATATCTAATCAAGAAACACATGGCAATGGSTGAGCGGTACCT 294
UY 273 Leu 273
|||
DB 295 CTG 297

Search completed: December 7, 2002, 15:44:08
Job time : 3068 secs

•
•
•

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_P2n model

Run on: December 7, 2002, 11:29:36 : Search time 262 seconds

(without alignments)
2793.509 Million cell updates/sec

Title: us-09-831-458a-12

Sequence: 1 MSISKEPRVQGLTIGTGH PRIVNWYVKKPACPEDE 325

Scoring table:

OLIGO
Xgapop 60.0, Xgapext 60.0
Ygapop 60.0, Ygapext 60.0
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 1125934159 residues

Word size: 5

Total number of hits satisfying chosen parameters: 241264

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame_P2n model -DBY-X1b
-O/cgna2_1/USPro/Spool/US0831458/runtat_05122002_103417_16781/FF_query.fasta_1_519
-DB_N_Geneseq_101002 -OFMT=lastap -SUFFIX=oli.rng -MINMATCH=0.1 -DOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TPANS=bimanda.oli
-LIST=45 -DOCCALIGN=200 -THR_SCORE=quality -THR_MIN=5 -ALTGEN=45 -MPER=LOCAL
-OUTFMT=pio -NORMEXT -HEAPEXT=500 -MINLEN=0 -MAXLEN=200000000
-USER=0509831458 -O3N 1 1 0 -runtat_05122002_103417_16781 -RCPU 6 -RCPU=3
-NO_X1XPY -NO_MMAP -IARGXUDFERY -NFG_STOPPS=0 -MAIT -FONQ=05 -PRV_T1MPROT=120
-MARR_T1MPROT=10 -IHRFAPS=1 -X1APop=60 -X1GAPEXT=60 -F3APop=6 -F3GAPEXT=7
-YGAOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_101002.*
1: /SIDS2/qcadata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/qcadata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/qcadata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/qcadata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/qcadata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/qcadata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/qcadata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/qcadata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/qcadata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/qcadata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/qcadata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/qcadata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/qcadata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/qcadata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/qcadata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/qcadata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/qcadata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/qcadata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/qcadata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/qcadata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/qcadata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/qcadata/geneseq/geneseq-emb1/NA2001B.DAT.*
23: /SIDS2/qcadata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/qcadata/geneseq/geneseq-emb1/NA2002.DAT.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Desc
1	325	100.0	1542	21	AAA27055	Human
2	168	51.7	1643	20	AAV84361	Human
3	168	51.7	1643	22	AAV15469	Human
4	168	51.7	1643	24	AAH41993	Human
5	97	29.8	792	22	AAH8239	Human
6	97	29.8	792	22	AAH16026	Human
7	84	25.8	553	21	AAH67709	Human
8	78	24.0	1212	21	AAH41224	Human
9	78	24.0	1215	21	AAH8740	Human
10	78	24.0	1215	21	AAH65383	Human
11	78	24.0	1312	14	AAU6648	Human
12	78	24.0	1312	22	AAH5461	Human
13	55	16.9	433	22	AAH99465	Human
14	51	15.7	592	22	AAH60904	Human
15	51	15.7	592	22	AAH28884	Human
16	51	15.7	592	22	AAH09194	Human
17	51	15.7	592	22	AAH35083	Human
18	51	15.7	592	22	AAH40799	Human
19	51	15.7	592	24	AAH9548	Human
20	50	15.4	152	22	AAH43421	Human
21	50	15.4	152	22	AAH48762	Human
22	50	15.4	152	22	AAH21864	Human
23	50	15.4	152	22	AAH48027	Human
24	50	15.4	152	22	AAH38856	Human
25	50	15.4	152	24	AAH21963	Human
26	43	13.2	1212	24	AAH41224	Human
27	35	10.8	1212	24	AAH41224	Human
28	29	9	265	20	AAH89716	Human
29	29	9	606	24	AAH20341	Human
30	30	9	1183	16	AAH6014	Human
31	31	9	1516	22	AAH42221	Human
32	32	9	2253	18	AAH2559	Human
33	33	9	2648	19	AAH2558	Human
34	34	9	2755	24	AAH12899	Human
35	35	9	2757	22	AAH41177	Human
36	36	9	2757	24	AAH49160	Human
37	37	9	2757	24	AAH40585	Human
38	38	9	2781	24	AAH49159	Human
39	39	9	2889	24	AAH49162	Human
40	40	9	2892	24	AAH49161	Human
41	41	9	3512	24	AAH49163	Human
42	42	9	4418	22	AAH2159	Human
43	43	9	16424	22	AAH68448	Human
44	44	9	16424	22	AAH2567	Human
45	45	9	14005	24	AAH98500	Human

ALIGNMENTS

RESULT 1

ID AAA27055 standard: cDNA: 1542 bp

XX AAA27055:

XX 22-AUG-2000 (first entry)

XX Human cell surface receptor protein cDNA sequence #12.

XX Human: HCSR: cytosolic; antiarthritic; antirheumatic;

XX immunosuppressive; antiarteriosclerotic; antibacterial;

XX neuroprotective; nontoxic; anticonvulsant; cancer; leuk

XX melanoma; rheumatoid arthritis; asthma; atherosclerosis

XX Alzheimer's diseases; multiple sclerosis; epilepsy; ss

XX Homo sapiens.

XX

FH KEY location/qualifiers
 FT CDS 269..1246
 FT /tag= d
 FT /product= "HCSR-12"
 PN W0209028032-A2.
 XX 18-MAY-2000.
 PD
 XX 12-MAY-1999: 99WO-US26742.
 XX 12-MAY-1999: 98US-0191280.
 PR 07-JUN-1998: 98US-0206437
 PR 08-MAR-1999: 99US-0123404.
 XX
 PA (INCY-1) INCYTE PHARM INC.
 XX
 PT Tanq YT, Corley NC, Guebler KJ, Yue H, Baughn MR, Lal P;
 PI Hillman JL, Bandman O, Azimzai Y, Au-Young J;
 XX
 DR WPI: 2900-37646/72.
 DR P-PSDB: AAY94345.
 XX
 PS New human cell surface receptor protein and polynucleotide useful for
 PT diagnosis, prevention and treatment of cancer, immune disorders,
 PT infection and neuronal disorders
 PT
 XX
 PS Claim 9; page 36-97; 97pp; English.
 XX
 CC The present sequence encodes a novel human cell surface receptor protein
 CC (HCSR-12) designated HCSR-12. The nucleotide sequence was identified in
 CC Incyte Clone 344986 from the cDNA library SPLN0909, which was made
 CC from RNA isolated from diseased spleen tissue. A number of Incyte clones
 CC were used to assemble the consensus sequence. BLAST analysis showed that
 CC the sequence is homologous to non-CD4 glycoprotein gp120 receptor
 CC AR3188. HCSR and its antagonist are useful for preventing or treating
 CC disorders associated with decreased or increased expression of activity
 CC of HCSR. Such disorders include cancers such as leukaemia and melanoma,
 CC immune disorders such as rheumatoid arthritis, asthma and
 CC atherosclerosis, bacterial and parasitic infections and neuronal
 CC disorders such as akathisia, Alzheimer's disease, multiple sclerosis and
 CC epilepsy. Polynucleotides encoding HCSRs may be used as hybridisation
 CC probes to diagnose these conditions. Anti-HCSR and antibodies may be used
 CC as antagonists, as a targeting or delivery mechanism for brining
 CC pharmaceutical agents into contact with cells or tissues expressing
 CC HCSR and for diagnosis of HCSR-related disorders. HCSR and its
 CC catalytic or immunogenic fragments are useful for drug screening using
 CC libraries of compounds.
 XX
 SU Sequence 1542 BP, 382 A, 417 C, 408 G, 435 T, 0 other.
 Alignment Scores:
 Pred. No.: 1,396-313 Length: 1542
 Score: 325.00 Matches: 325
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0
 US-09-831-458a-12 (1-345) x AAY27055 (1 1542)
 QY 1 MetSerAspSerLysGluProArValGlnGlnLeuGlyLeuLeuGlyCysLeuGlnHis 20
 DB 259 ATGAGTGAATCTCAAGCAACCAATGATGACCAATGAGGCTCTGAGGCTTGGTGGAT 328
 QY 21 GVALAGValLeuGlnLeuGlnSerPheMetLeuLeuValGlnValLeuValAlaIle 40
 DB 329 GAGGAGCTGATGATGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 498
 QY 41 LeuValGlnValSerLysValProSerSerLeuSerGlnGlnGlnSerGlnGlnAspAla 60
 DB 389 GTTGTAAAGTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 448

QY 61 TCTCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 80
 DB 449 ATCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 508
 QY 81 TCTCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 100
 DB 509 CTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 568
 QY 101 TCTCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 120
 DB 569 AATCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 628
 QY 121 TCTCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 140
 DB 629 TCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 688
 QY 141 TCTCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 160
 DB 689 GTCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 748
 QY 161 TCTCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 180
 DB 749 AATCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 808
 QY 181 TCTCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 200
 DB 809 AATCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 868
 QY 201 TCTCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 220
 DB 869 GTCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 928
 QY 221 TCTCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 240
 DB 929 GTCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 988
 QY 241 TCTCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 260
 DB 989 GTCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 1048
 QY 261 TCTCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 280
 DB 1049 GTCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 1108
 QY 281 TCTCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 300
 DB 1109 GTCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 1168
 QY 301 TCTCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 320
 DB 1169 GTCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 1228
 QY 321 TCTCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 340
 DB 1229 GTCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 1288
 RESULT 2
 ID AAY84361
 XX AAY84361 standard; cDNA to mRNA; 1643 BP.
 AC AAY84361:
 XX
 DT 30-MAR-1999 (first entry)
 XX
 DE Human liver cDNA clone H901347.
 XX
 KM Transmembrane protein; H901347; human; lectin; receptor; liver; ds.
 OS Homo sapiens.
 XX
 XX
 FH KEY location/qualifiers
 FT CDS 25..915
 FT /tag= d

CC necessary for initiating an immune response. DC-SIGN is also able
 CC to bind to the HIV envelope protein gp120 and to facilitate HIV-1
 CC entry into DC. The cDNA was obtained by RT-PCR amplification (see
 CC AAA88741-42) of DC RNA. The invention relates to the use of a
 CC compound that binds to a C-type lectin (especially DC-SIGN) on the
 CC surface of a DC, and in the preparation of a composition for
 CC modulating, especially inducing, the immune response in an animal,
 CC in particular a human or other mammal. The composition modulates
 CC the interactions between a DC and a T-cell, more specifically
 CC between a C-type lectin on the surface of a DC and an ICAM receptor
 CC on the surface of a T-cell. It is used to prevent/inhibit immune
 CC responses to specific antigens, for inducing tolerance, for
 CC immunotherapy, for immunosuppression, for the treatment of
 CC autoimmune diseases, the treatment of allergy, and/or for
 CC inhibiting HIV infection (claimed). The compound that binds to
 CC the C-type lectin is chosen from mannose, fucose, plant lectins,
 CC antibodies, sugars, proteins and antibodies against C-type lectins.
 XX

Sequence 1215 BP; 315 A, 321 C, 349 G, 230 T, 0 other,

Alignment Scores:
 Pred No. 6, 83e-68 Length: 1215
 Score: 78.00 Matches: 78
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 24.00% Indels: 0
 DB: Caps: 0

US-09-831-458A-12 (1-325) x AAA88740 (1-1215)

QY 58 GlnAspAlaIleTyrGlnAsnLeuThrGlnLeuLysAlaIleValCylGluLeuSerGlu 77
 Db 220 TAAATACCAATATTAATTAAGCTTGAATCAATATTAATTAATTAATTAATTAAG 279
 QY 78 LysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 97
 Db 230 AATATCAATATTAATTAAGCTTGAATCAATATTAATTAATTAATTAATTAAG 279
 QY 98 LeuProGlnLysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 117
 Db 340 CTTCACATTAATTAAGCTTGAATCAATATTAATTAATTAATTAATTAAG 399
 QY 118 ValGlyGlnLeuProGlnLysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGln 135
 Db 400 GTTACATTAATTAAGCTTGAATCAATATTAATTAATTAATTAATTAAG 453

RESULT 10

AA065383 ID AAC65383 standard; cDNA; 1215 BP.

AC AAC65383;

DT 12-FEB-2001 (first entry)

DE Human C-type lectin receptor DC-SIGN nucleotide sequence.

XX Human; immunomodulatory; antiallergic; anti-HIV; antiinflammatory;

XX antidiabetic; antihypertensive; antirheumatic; antidiarrhoeal; vaccine;

XX C-type lectin receptor; immunotherapy; immunosuppression;

XX transplant rejection; autoimmune disease; thyroiditis;

XX rheumatoid arthritis; multiple sclerosis; autoimmune diabetes;

XX systemic lupus erythematosus; HIV infection; allergy; DC-SIGN, ss.

XX Homo sapiens.

OS

PN WO200063251 A1.

XX

XX 26-SEP-2000

XX 19-APR-2000; 2000MO-NL00253.

XX 19-APR-1999; 99EP-0201204.

XX 20-JAN-2000; 2000US 0176221.

XX (UNVI-) UNIV NIJMEGEN.

PA Finder CC, Cejicbeck THH, Van Kooyk Y, Torensmas R;

XX WU; 2000-050424/63.

XX F-ESTAB; AAR2914.

XX Methods for modulating dendritic cell and T cell interaction by using a

XX compound, such as mannose carbohydrates, that binds to a C-type lectin

XX on a dendritic cell surface, useful in the treatment of autoimmune

XX diseases and allergy.

XX Claim 11, Fig 9, 66pp; English.

XX The present sequence encodes DC-SIGN, a 44kDa C-type lectin receptor. New

XX methods for modulating dendritic cell and T cell interaction are

XX disclosed. The methods involve using a compound that binds to a C-type

XX lectin on the surface of a dendritic cell. The methods are useful for

XX preventing or inhibiting immune responses to specific antigens, for

XX inducing tolerance, for immunotherapy, for immunosuppression (i.e. for

XX preventing transplant rejection) for the treatment of autoimmune

XX diseases (e.g. thyroiditis, rheumatoid arthritis, multiple sclerosis,

XX autoimmune diabetes, systemic lupus erythematosus), HIV infection and/or

XX for the treatment of allergy. Antibodies are useful for detecting the

XX presence of dendritic cells in a biological sample and for determining

XX the presence and/or expression of the C-type lectins or their fragments

XX or epitopes in a biological sample.

XX Sequence 1215 BP; 315 A, 321 C, 349 G, 230 T, 0 other,

XX

XX Alignment Scores:

XX Pred. No. 6, 83e-68 Length: 1215

XX Score: 78.00 Matches: 78

XX Percent Similarity: 100.00% Conservative: 0

XX Best Local Similarity: 100.00% Mismatches: 0

XX Query Match: 24.00% Indels: 0

XX DB: Caps: 0

XX US-09-831-458A-12 (1-325) x AAC65383 (1-1215)

QY 58 GlnAspAlaIleTyrGlnAsnLeuThrGlnLeuLysAlaIleValCylGluLeuSerGlu 77

Db 220 TAAATACCAATATTAATTAAGCTTGAATCAATATTAATTAATTAATTAATTAAG 279

QY 78 LysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 97

Db 230 AATATCAATATTAATTAAGCTTGAATCAATATTAATTAATTAATTAATTAAG 279

QY 98 LeuProGlnLysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 117

Db 340 CTTCACATTAATTAAGCTTGAATCAATATTAATTAATTAATTAATTAAG 399

QY 118 ValGlyGlnLeuProGlnLysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGln 135

Db 400 GTTACATTAATTAAGCTTGAATCAATATTAATTAATTAATTAATTAAG 453

RESULT 11

AA036648 ID AAC66648 standard; cDNA; 1312 BP.

AC AAC66648;

DT 08-JUN-1993 (first entry)

DE Sequence encoding non-CD4 glycoprotein gp120 receptor protein.

XX Surface CD4 protein, glycoprotein receptor protein; gp120 receptor,

XX HIV; diagnosis; therapy; ss.

XX Homo sapiens.

OS

PN

XX Key Location/variants

XX	AAH9365:			
XX	16-01-2001 (first entry)			
XX	Human protein encoding cDNA sequence. SEQ ID NO:200.			
DE	Human: cancer; ulcer; HIV infection; human immunodeficiency virus;			
XX	antitumour; antirheumatic; arthritis; immunosuppression;			
XX	antibacterial; endocrine; cardiac; central nervous system; vitreous;			
XX	anti-HIV; fungicide; anti-invasive; cardiovascular; osteoporosis;			
XX	antibacterial; antineoplastic; antitumor; antidiabetic; eczema;			
XX	dermatological; antiallergic; antidiabetic; antidiabetic; cytotoxic;			
XX	neuroprotective; antidepressant; osteoporosis; immunomodulatory;			
XX	immunostimulant; gene therapy; antineoplastic; vaccine; inflammation;			
XX	antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;			
XX	cardiac dysfunction; neuropathology; cardiac myopathy; antineoplastic;			
XX	specific disease; haematopoietic disorder; platelet disorder; asthma;			
XX	thrombocytopenia; osteoporosis; severe combined immunodeficiency;			
XX	allergic rhinitis; diabetes; multiple sclerosis; depression;			
XX	Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;			
XX	neurological disorder; ss.			
XX	Hemo sapientis.			
XX	W0200153455-A2.			
XX	26-01-2001.			
XX	22-DEC 2000; 2000WO/0835017.			
XX	23-DEC-1999; 99US-0471275.			
XX	21-JAN-2000; 2000US-0488725.			
XX	25-APR-2000; 2000US-0552317.			
XX	(HYSE-) HYSEQ INC.			
XX	Tang YT, Liu C, Dymnac RT;			
XX	WPI: 2001/457603/49.			
XX	P-PSDB; AAM25424.			
XX	Isolated human polynucleotides encoding polypeptides, useful for the			
XX	treatment and diagnosis of e.g. cancer, ulcers and HIV infection.			
XX	Claim 1: Page 387: 1217pp; English			
XX	AAH9456 to AAH9904 encode the human proteins given in AAM25225 to			
XX	AAM25953. The proteins can have activities based on the tissues and			
XX	cells they are expressed in, such as: antineoplastic; antirheumatic;			
XX	antibacterial; immunosuppressive; antibacterial; endocrine; cardiac;			
XX	central nervous system; viricide; anti-HIV; fungicide; antitumor;			
XX	cardiovascular; antineoplastic; antineoplastic; haematologic; vaccine;			
XX	antitumor; osteoporosis; dermatological; antiallergic; antineoplastic;			
XX	antidiabetic; cytotoxic; neuroprotective; antidepressant; osteoporosis;			
XX	antiparkinsonian; and immunostimulant. The proteins and polynucleotides			
XX	encoding them can be used in gene therapy, antisense therapy and vaccine			
XX	production. The proteins and polynucleotides are useful for screening for			
XX	agonists or antagonists of a protein and for the treatment and diagnosis			
XX	of disorders associated with the activity of a protein e.g. inflammation,			
XX	rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,			
XX	neuropathology, cardiac myopathy, viral, bacterial, HIV and fungal			
XX	infections, autoimmunity, multiple sclerosis, haematological disorders,			
XX	anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,			
XX	osteoporosis, severe combined immunodeficiency, eczema, allergic			
XX	rinitis, asthma, diabetes, cancer, multiple sclerosis, depression,			
XX	Alzheimer's disease, Parkinson's disease, neurodegenerative and			
XX	neurological disorders.			
XX	Sequence 433 RP: 110 A; 114 G; 128 G; 91 T; 0 other;			
XX	Alignment Scores: 1,999-45 Length: 443			
XX	Entry No. 1,999-45 Length: 443			

Percent Similarity:	55.00	Matches:	55
Best Local Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	16.928	Indels:	0
DB:	22	Gaps:	0

US-09-831-458A-12 (1-325) x AAH99365 (1-433)

```

C7      59 cttggggtattatcttgcacacccctttccttccttgagaaataaaatavtactgtaaccscgch 77
        |||
E6      216 caaacacgaacatcaaacagaacatgacacaaatcaaaaattcaattacaagcagaaatcaaaag 275
        |||
WY      78 lysesltskndnclnclatcttgtctctctctctctctctctctctctctctctctctctctctct 97
        |||
I6      276 aaatcgaacctgcgaacacgaatcaaacacgaatcaaacacgaatcaaacacgaatcaaacacgaat 335
        |||
G7      98 tcaattcttgagsgcatltskndnclnclatcttgtctctctctctctctctctctctctctct 112
        |||
I6      336 cttaaagaaaaatcaaatcttcaaacacgaatcaaacacgaatcaaacacgaatcaaacacgaat 380
        |||
RESULT 14
ID      ABA60904 standard; DNA; 592 BP.
XX
AC      ABA60904:
XX
DT      01-FEB-2002 (first entry)
XX
DE      Human foetal liver single exon nucleotide acid probe #9209.
XX
KM      Human: foetal liver gene expression; single exon nucleotide acid probe; ss.
CS      Homo sapiens.
FN      W0200157277-A2.
PD      09-AUG-2001.
XX
XX      30-JAN-2001; 2001wo-US00669.
XX
PF      04-FEB-2000; 200000S-018031Z.
PR      26-MAY-2000; 200000S-020745Z.
PP      30-JUN-2000; 200000S-060840R.
PI      03-AUG-2000; 200000S-063236P.
PR      21-SEP-2000; 200000S-023468T.
PP      27-SEP-2000; 200000S-023635Y.
PE      04-OCT-2000; 200000S-0024261.
XX
EA      (MELI.) MOLETO/LAB DYNAMICS INC.
XX
ET      Peun SS; Hardest UK; Chou W.; Rauh EG;
XX
DR      WPT, 2001 483447/752.
XX
TT      Human genome derived single exon nucleotide acid probes useful for
        analyzing gene expression in human foetal liver -
XX
PS      Claim 1: seq ID NO 9209; 63bp + sequence listing; English.
XX
RE      The invention relates to a single exon nucleotide acid probe for
        measuring human gene expression in a sample derived from human foetal
        liver. The single exon nucleotide acid probes may be used for predicting,
        monitoring and displaying gene expression in samples derived from human
        foetal liver. The present sequence is a single exon nucleotide acid
        probe of the invention.
CC      Note: The sequence data for this patent did not form part of the
        printed specification, but was obtained in electronic format directly
        from WIPO at http://wipo.int/pub/publicated\_pat\_sequences.
XX
S2      Sequence C92 BP: 154 A; 149 G; 168 G; 121 T; 6 other;
    
```

Alignment Scores:
Pred. No.: 2,560-41 Length: 592

KM gene to disorder: SS
 XX Homo sapiens
 XX W0200157272-A2.
 XX
 XX 09-AUG-2001.
 XX
 XX 30-JAN-2001: 2001WO-0500663.
 XX
 XX 04-FEB-2000: 2000US-0180312.
 XX 26-MAY-2000: 2000US-0207456.
 XX 30-JUN-2000: 2000US-0608408.
 XX 23-AUG-2000: 2000US-0632366.
 XX 21-SEP-2000: 2000US-0234687.
 XX 27-SEP-2000: 2000US-0234686.
 XX 04-OCT-2000: 2000US-0024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 XX Penn St. Hanzel FR. Chen W. Park DR.
 XX WPI: 2001-488897/53.
 XX
 XX Human genome derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human placenta -
 XX
 XX Claim 25: SEQ ID NO 22542; 654pp; English.
 XX
 XX The present invention relates to single exon nucleic acid probes (SENP).
 XX The present sequence is one such probe. The probes are useful for
 XX producing a microarray for predicting, measuring and displaying gene
 XX expression in samples derived from human placenta. The probes are useful
 XX for antenatal diagnosis of human genetic disorders.
 XX
 XX Sequence 152 HP, 37 A; 48 C, 38 G, 23 T, 0 other:
 XX
 XX Alignment Scores:
 XX Pref. No.: 7e-41 Length: 152
 XX Score: 60.00 Matches: 50
 XX Percent Similarity: 100.00% Conservative: 3
 XX Host Local Similarity: 100.00% Mismatch: 0
 XX Query Match: 15,388 Indels: 0
 XX DB: 22 Gaps: 0
 XX
 XX US-09-831-458A-12 (1-325) x AA153856 (1-152)
 XX
 XX 189 ArgLacGysArgHisCysProGlySerThrPheGluGluGlyAsnGlySerThrMet 208
 XX 3 CGCGTGGTGGTAACTGGTAAAGGAGTGAATATGTCGAAAGCAAATGTTACTTCATG 62
 XX
 XX 209 SerAsnSerLysArgAsnTrpHisAspSerValIleAlaGlyGlnGluValATGAlaGln 238
 XX 63 TCTAATCCGACGGGTAACATGACATGCGCGCAATGATGCTGCAAGAGGAGGCTGAA 122
 XX
 XX 229 LeuValValIleLeuThrAlaGluGlu 238
 XX 123 GTGCTGTATATCAAAATGCTGAGGAGCA 152
 XX
 XX RESULT 25
 XX AHS21963
 XX AHS21963 standard. DNA; 152 BP.
 XX
 XX AHS21963;
 XX 19-AUG-2001 (first entry)
 XX
 XX Human genome-derived single exon probe (HP) from lung SEQ ID NO 21954.
 XX
 XX Human ds; single exon probe; asthma, lung cancer, COPD, ILD;
 XX chronic obstructive pulmonary disease; interstitial lung disease;
 XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
 XX tuberculous sclerosis, Gaucher's disease, Niemann-Pick disease;

KM Hirschman-Podlak syndrome, sarcoidosis, pulmonary haemosiderosis;
 KM pulmonary histiocytosis, lymphangioleiomyomatosis, Karagane syndrome;
 KM pulmonary alveolar proteinosis, fibrocystic pulmonary dysplasia;
 KM primary ciliary dyskinesia, pulmonary hypertension,
 KM hyaline membrane disease, open reading frame, ORF.
 XX
 XX Homo sapiens.
 XX
 XX W020016004-A2.
 XX
 XX 15 DEC 2001.
 XX
 XX 30 JAN 2001: 2001WO-0500663.
 XX
 XX 04-FEB-2000: 2000US-180312P.
 XX 26-MAY-2000: 2000US-0207456P.
 XX 30-JUN-2000: 2000US-0608408.
 XX 23-AUG-2000: 2000US-0632366.
 XX 21-SEP-2000: 2000US-0234687P.
 XX 27-SEP-2000: 2000US-0234686P.
 XX 04-OCT-2000: 2000US-0024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 XX Penn St. Hanzel FR. Chen W. Park DR.
 XX WPI: 2002-114183/15.
 XX
 XX Spatially-addressable set of single exon nucleic acid probes, used to
 XX measure gene expression in human lung samples -
 XX
 XX Claim 4: SEQ ID NO 21954; 654pp; English.
 XX
 XX The invention relates to a spatially-addressable set of single exon
 XX nucleic acid probes for measuring gene expression in a sample derived
 XX from human lung comprising single exon nucleic acid probes having one of
 XX 12614 nucleic acid sequences mentioned in the specification, or their
 XX complements or the 12387 open reading frames derived from the 12614
 XX probes. Also included are a microarray comprising the novel set of
 XX probes; the novel set of probes which hybridize at high stringency to a
 XX nucleic acid expressed in the human lung; measuring gene expression in a
 XX sample derived from human lung, comprising (a) contacting the array with
 XX a collection of detectably labeled nucleic acids derived from human lung
 XX mRNA; and (b) measuring the label detectably bound to each probe of
 XX the array; identifying exons in a eukaryotic genome, comprising
 XX (a) algorithmically predicting at least one exon from genomic sequences
 XX of the eukaryote; and (b) detecting specific hybridisation of detectably
 XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 XX having a fragment identical to the predicted exon, the probe is included
 XX in the above mentioned microarray; assigning exons to a single gene,
 XX comprising (a) identifying exons from genomic sequence by the method
 XX above and (b) measuring the expression of each of the exons in several
 XX tissues and/or cell types using hybridisation to a single exon
 XX microarrays having a probe with the exon, where a common pattern of
 XX expression of the exons in the tissues and/or cell types indicates that
 XX the exons should be assigned to a single gene; a peptide comprising one
 XX of 12011 sequences mentioned in the specification, or encoded by the
 XX protein; a set of reading frames (ORF). The probes are used for gene
 XX expression analysis, and for identifying exons in a gene, particularly
 XX using human lung derived mRNA and for the study of lung diseases
 XX such as asthma, lung cancer, chronic obstructive pulmonary disease
 XX (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 XX fibrosis, neurofibromatosis, tuberculous sclerosis, Gaucher's disease,
 XX Hirschman-Podlak syndrome, Niemann-Pick syndrome, sarcoidosis, pulmonary
 XX haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 XX pulmonary alveolar proteinosis, Karagane syndrome, fibrocystic
 XX pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 XX and hyaline membrane disease. The present sequence is a single exon
 XX probe open reading frame of the invention.
 XX Note: the sequence data for this patent did not form part
 XX of the printed specification, but was obtained in electronic
 XX format directly from WPI at
 XX ftp.wpi.int/pub/published_prt_sequences.

XX The present sequence is that of the coding region of the gorilla
 CC DC-SIGN gene. DC-SIGN is expressed on dendritic cells and is known
 CC to provide a mechanism for transport of HIV-1 virus to the lymph
 CC nodes. HIV-1 binds to the extracellular portion of DC-SIGN and
 CC infects the unfertilized T cells in the lymph nodes via their
 CC CD4 proteins. This ultimately leads to compromise of the immune
 CC system and to full-blown AIDS. The invention comprises a
 CC comparative genomics approach to identify specific gene changes
 CC responsible for differences in functions and diseases distinguishing
 CC humans from non-humans, particularly primates including gorilla,
 CC orangutan and especially chimpanzee. The evolutionary significance
 CC of a nucleotide change is determined by the ratio of the
 CC non-synonymous substitution rate (KA) to the synonymous rate (KS)
 CC of the nucleotide sequence. Polynucleotide and polypeptide
 CC sequences corresponding to evolved traits may be relevant to human
 CC diseases or conditions such as unique or enhanced human brain
 CC functions, longer human life spans, susceptibility or resistance to
 CC disease, including AIDS and cancer, and aesthetic traits such as
 CC hair growth. Ka/Ks ratios for chimpanzee/human, human/gorilla and
 CC chimpanzee/gorilla DC-SIGN cDNA sequences are 1.3, 0.87 and 1.3,
 CC respectively. It is theorized that chimpanzee resistance to
 CC progression to full-blown AIDS may be due in part to inability of
 CC HIV-1 to bind to chimpanzee DC-SIGN for transport to the lymph
 CC nodes. After determining the 3-dimensional structure of DC-SIGN,
 CC a rational drug design approach can be used to mimic the effects
 CC of chimpanzee DC-SIGN without interfering with the normal functions
 CC of human DC-SIGN.

XX Sequence 1212 BP; 315 A; 319 C; 349 G; 229 T; 0 other;

Alignment Scores:

Prod. No.: 4.19e-25
 Score: 35.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 10.77%
 DB: 24
 Gaps: 0

US-09-831-458a-12 (1-325) x ABA91225 (1-1212)

Q7 59 LysAlaAlaValAlcylGluLeuSerGluLysSerLysLeuGluGluLeu 88
 DB 253 AAAGCTGCTACTGAGTGAAGCTCTAGCAAAATGCAATGATATAGGAGCTG 412

QY 89 ThcGlnLeuLysAlaAlaValAlcylGluLeuFrogLcysSerLys 103
 |||
 DB 313 ACCCACTCTGAAGGCTGTGAGCTGACCTTCTAGACAAATCTAAG 357

RESULT 28

AAV89716
 ID AAV89716 standard; cDNA; 265 BP.

XX AAV89716;

XX 15-FEB-1999 (first entry)

XX EST clone CF785.

XX Human; secreted protein; expressed sequence tag; EST; haematopoiesis;
 KW tissue growth; actinin; tubulin; chemokinesis; haemostatic;
 KW receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti tumour;
 KW gene therapy; ss.

XX Homo sapiens.

XX WO9845436-A2

XX 15-OCT-1998.

XX 10-APR-1998; 98MO-US06955.

XX 10-APR-1997; 97US-0838821

XX (GEMV) GENETICS INST INC.

XX Ag-sting M1, Mark's F, Javallier-Et, Mary JM, Methylene D;

XX Racie LA, Spaulding V, Treacy M;

XX WPI; 1999-070077/06.

XX New polynucleotides encoding human secreted proteins - derived from
 PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
 PT ovary, pituitary, retina and colon cDNA libraries.

XX Claim 1; Page 367; 618pp; English.

XX The present sequence represents a human expressed sequence tag (EST).
 CC The polynucleotide, which is a secreted EST, and the encoded protein
 CC are predicted to have useful biological activities which would make
 CC them suitable for treating, preventing or ameliorating medical
 CC conditions in humans and animals, although no supporting data is
 CC given. Suggested activities include nutritional activity, immune
 CC stimulating or suppressing activity, haematopoiesis regulating
 CC activity, tissue growth activity, activin/inhibin activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic
 CC activity, anti-thrombotic activity, anti-inflammatory activity,
 CC cadherin/tumour invasion suppressor activity, tumour inhibition
 CC activity. The polynucleotide may also be useful for gene therapy.

XX Sequence 265 BP; 74 A; 54 C; 74 G; 74 T; 0 other;

Alignment Scores:

Prod. No.: 7.45
 Score: 9.00
 Percent Similarity: 100.00%
 Best Local Similarity: 2.77%
 Query Match: 20
 DB: 20
 Gaps: 0

US-09-831-458a-12 (1-325) x AAV89716 (1-265)

QY 34 AlaGlyValLeuValAlaLeuVal 42

DB 176 GCTGAGCTGTTGCTGAGTGAAGCTCTAGCAAAATGCAATGATATAGGAGCTG 202

RESULT 29

ABR30341
 ID ABR30341 standard; cDNA; 606 BP.

XX ABR30341;

XX 23-APR-2002 (first entry)

XX Human G-protein-coupled protease #111.

XX Human; ss; gene; G-protein-coupled protease; gene therapy;
 KW transgenic; protease mediated disorder; proliferative disorder;
 KW differential disorder; developmental disorder;
 KW haematopoietic disorder;

XX Homo sapiens.

XX US6331427-B1.

XX 18-DEC-2001.

XX 26-MAR-1999; 99US-0280116.

XX 26-MAR 1999; 99US-0280116.

XX (MILL-) MILLENNIUM PHARM INC.

XX Robison KE;

XX WPI; 2002-129545/17.

XX New polynucleotides encoding protease homologs of the G-protein-coupled
PT protease family, useful in identifying agonists and antagonists for
PT diagnosis and treatment of protease mediated disorders -
XX
PS Disclosure: Column 155-156; 246pp; English.

XX The invention relates to an isolated human protease nucleic acid molecule
CC comprising a nucleotide sequence of 546 base pairs, one of 266 fully
CC defined in the specification. Also disclosed are production of an
CC isolated polypeptide encoded by the nucleic acid, comprising introducing
CC the nucleic acid into a host cell and culturing under conditions to
CC express the protein from the nucleic acid, use of an antibody to
CC detect the encoded protein in a sample and to modulate its *in vivo*
CC activity, identifying agents that bind to the protein and identification
CC of a polynucleotide agent that modulates the expression of the nucleic
CC acid or its complement (i.e. gene therapy). The nucleic acid can be used
CC to identify an agent that modulates the expression or activity of the
CC nucleic acid, and can be used to isolate the protein. The nucleic acid
CC can be used in diagnostic assays for determining nucleic acid expression
CC as well as activity in the context of a biological sample (e.g., blood,
CC serum, cells, tissue) to determine whether an individual has a disease or
CC disorder, or is at risk of developing a disease or disorder, associated
CC with aberrant expression or activity of the nucleic acid. The nucleic
CC acid can be used to detect mutations in protease genes and gene
CC expression products such as mRNA. The nucleic acid can be used as
CC hybridisation probes to detect naturally-occurring genetic mutations in
CC a protease gene. The nucleic acid can be used in drug screening methods
CC to identify agonists and antagonists that can be used to diagnose and
CC treat such protease mediated disorders e.g., proliferative,
CC degenerative, developmental or hematopoietic disorders. The nucleic
CC acid can be used as probes, primers, in biological assays, to determine
CC patterns of gene expression, to design ribozymes and to construct
CC transgenic animals. The present sequence represents one of the 268
CC disclosed human G-protein-coupled protease cDNA sequences.

XX
SQ Sequence 606 BP; 129 A; 192 C; 165 G; 115 T; 5 other;

Alignment Scores:
Pred. No. 16.3 Length: 606
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.77% Indels: 0
DB: 24 Gaps: 0

US-09-831-458a-12 (1-325) x ABK30341 (1-606)

QY 34 AAGGYYVallenuValAlaIleLeuVal 42
|||||
Db 199 GCTGAGCTTTCTCTGCGCATCTTG57G 225

RESULT 30
AAT86014
ID AAT86014 standard; cDNA: 1183 BP.
AC AAT86014;
XX
DT 08-MAY-1998 (first entry)
XX
DE Human beta mellitin cDNA fragment 2.
XX
KW Mellitin: myotube formation; muscle tissue; therapy; membrane protein;
KW myoblasts; adhesion; fusion; medicinal; muscle disorder; human; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1183
FT /tag= a
FT /product= beta mellitin
FT /note= "partial coding sequence"
FT /codon_start= 2

XX
PN W09731109-A1.
XX
PD 28-AUG-1997.
XX
PF 17-OCT-1996; 96W0-JP03017.
XX
PK 23-FEB-1996; 96JP-U061756.
XX
PA (MOCH) MOCHIDA PHARM CO LTD.
XX
FI Fujisawa A, Mizushima G, Oikawa N, Shirakawa K, Yanaka
XX
DR WPI: 1997-435161/40.
XX
UK P-IPDB: AAW5722.
XX
PT Mellitin proteins involved with myotube formation in mice
XX
PS Claim 5; FIG 17; 262pp; Japanese.
XX
CC AAT86013 and AAT86014 encode fragments of a new human, novel
CC beta mellitin, which participates in the adhesion and fusion
CC in the course of myotube formation in muscle tissue. Mellitin
CC partial peptides can be used in the study of myotube form-
CC intracellular domain, a membrane-penetrating domain, a
CC domain, a disintegrin domain, a metalloproteinase domain,
CC domain and a cysteine rich domain although these are not
CC specification. Mellitin antagonists can be used as bone ad-
CC inhibitors, e.g. for the treatment of osteoporosis and for
CC and for prevention of cancer cell infiltration.

XX
SQ Sequence 1183 BP; 264 A; 419 C; 469 G; 241 T; 0 other;

Alignment Scores:
Pred. No. 40.6 Length: 1183
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.77% Indels: 0
DB: 18 Gaps: 0

US-09-831-458a-12 (1-325) x AAT86014 (1-1183)

QY 34 AAGGYYVallenuValAlaIleLeuVal 42
|||||
Db 998 GCTGAGCTTTGTTGCGCATCTTG67G 1024

RESULT 31
AAD03221
ID AAD03221 standard; cDNA: 1576 BP.
AC AAD03221;
XX
DT 13-JUN-2001 (first entry)
XX
DE Mouse type II integral membrane protein, cDNA.
XX
KW Mouse; type II integral membrane protein; cDNA; partial;
KW immunosuppressive; autoimmune disorder; allergy; viral in-
KW antigen presenting cell; bacterial infection; macrophage;
KW DC; ss.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 92..808
FT /tag= a
FT /product= "Mouse type II integral membra-
XX W0200119869-A1.

KP ovarian; ss.
 XX Homo sapiens.
 XX W020007423.A2.
 XX 07-DEC-2000.
 XX 25-MAY-2000: 2000MO-US14308.
 XX 27-MAY-1999: 99US-0136388.
 XX 09-JUN-1999: 99US-0142940.
 XX 28-JAN-2000: 2000US-0178717.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Ruben SM, Wei P, Ni J, Hastings CA, Shi Y;
 XX WPI: 2001-016507/02.
 XX Seven nucleic acid molecules encoding ADAM polypeptides containing a
 XX disintegrin and metalloprotease domain, useful in the prevention,
 XX treatment and diagnosis of cancer, immune disorders, cardiovascular
 XX disorders and neurological diseases -
 XX Claim 1: Page 264-265; 287pp; English.
 XX The present invention relates to seven members of the ADAMs (proteins
 XX which contain A Disintegrin And Metalloprotease domain) protein family.
 XX The ADAMs proteins and DNA may be used to treat disease, as a food
 XX additive or preservative, for chromosome identification, as probes
 XX for diagnosing a disorder related to the female reproductive system,
 XX particularly breast and/or ovarian cancer. They are also useful in the
 XX gene therapy of breast and ovarian cancer.
 XX Sequence 2757 BP: 640 A: 803 C: 787 G: 527 T: 0 other;
 XX
 XX Alignment Scores:
 XX Pred No: 67.9 Length: 2757
 XX Score: 9.00 Matches: 9
 XX Percent Similarity: 100.00% Conservative: 0
 XX Best Local Similarity: 100.00% Mismatches: 0
 XX Query Match: 2.77% Gaps: 0
 XX DB: 22 Gaps: 0
 XX US-09-831-458A-12 (1-325) x AAC91177 (1-2757)
 XX
 XX QY 34 AlaglyValleuValAlaIleleVal 42
 XX ||||||||||||||||||||||||||||
 XX DB 2113 GCTGAGTGTGGTGGCCATCTTGATG 2139
 XX
 XX RESULT 36
 XX AAD39160
 XX ID AAD39160 standard; cDNA; 2757 BP.
 XX AC AAD39160;
 XX XX 04-OCT-2002 (first entry)
 XX XX Human novel protease cDNA #2.
 XX XX Human novel human protein; NHP: protease; metelin-beta; ADAM 19;
 XX XX metalloprotease; drug screening; clinical trial methodology; evaluation
 XX XX Pharmacological, nutraceutical, gene therapy; chromosome 21 gene; ss.
 XX XX Homo sapiens
 XX XX Key Location/Qualifiers
 XX XX CDS 1..2757
 XX XX /tag= a
 XX XX /product= "Human novel protease #2"
 XX XX W0200236759-A2.

XX
 XX 10-MAY-2002.
 XX 30-OCT-2001: 2001WO-0548498.
 XX 01-NOV-2000: 2000US-244939P.
 XX (LEXI-) LEXICON GENETICS INC.
 XX Walke DW, Wilanowski NJ, Friddle CJ;
 XX WPI: 2002-479761/51.
 XX F: FSDB; AAF24351.
 XX New nucleic acid encoding a protein that shares sequence similarity
 XX with mammalian metelin-beta/ADAM 19 homolog metalloprotease for use in
 XX diagnosis and treatment, and in drug screening -
 XX Disclosure, Page 37-38, 48pp, English.
 XX
 XX The invention relates to novel human proteins (NHP), novel human
 XX proteases that shares sequence similarity with mammalian metelin-beta/
 XX ADAM 19 homolog metalloprotease, and their corresponding nucleic acid
 XX sequences. NHP DNA is useful for diagnosis, drug screening, clinical
 XX trial monitoring, treatment of diseases and disorders and
 XX pharmacological, cosmetic and nutraceutical applications. It is also
 XX useful as hybridisation probes for screening libraries, and assessing
 XX gene expression patterns (particularly using a microarray or high-
 XX throughput chip format). It is also useful in restriction fragment
 XX length polymorphism (RFLP) analysis to identify specific individuals,
 XX and as a probe to identify the corresponding mutant allele in the
 XX libraries. It is also useful for the detection of mutant
 XX NHPs or inappropriately expressed NHPs for the diagnosis of disease.
 XX It is also useful in gene therapy. The present sequence is novel human
 XX protease cDNA. Human NHP gene is located at chromosome 5.
 XX Sequence 2757 BP: 641 A: 802 C: 787 G: 527 T: 0 other;
 XX
 XX Alignment Scores:
 XX Pred No: 67.9 Length: 2757
 XX Score: 9.00 Matches: 9
 XX Percent Similarity: 100.00% Conservative: 0
 XX Best Local Similarity: 100.00% Mismatches: 0
 XX Query Match: 2.77% Gaps: 0
 XX DB: 24 Gaps: 0
 XX US-09-831-458A-12 (1-325) x AAD39160 (1-2757)
 XX
 XX QY 34 AlaglyValleuValAlaIleleVal 42
 XX ||||||||||||||||||||||||||||
 XX DB 2113 GCTGAGTGTGGTGGCCATCTTGATG 2139
 XX
 XX RESULT 37
 XX AAD30585
 XX ID AAD30585 standard; cDNA; 2757 BP.
 XX AC AAD30585;
 XX XX 21 MAY-2002 (first entry)
 XX XX Human protease, PKTS-18 cDNA.
 XX XX Human protease; PKTS-18 cDNA.
 XX XX Human protease; PKTS-18 cDNA.
 XX XX gastrointestinal disorder; autoimmune; inflammatory; cell proliferative;
 XX XX cardiovascular; developmental; epithelial; neurological; reproductive;
 XX XX AIDS; Acquired Immune Deficiency Syndrome; allergy; rheumatoid arthritis;
 XX XX anaemia; asthma; atherosclerosis; hypertension; myocardial infarction;
 XX XX hepatitis; cancer; psoriasis; Cushing's syndrome; hypothyroidism; eczema;
 XX XX epilepsy; Alzheimer's disease; Huntington's disease; Parkinson's disease;
 XX XX Pick's disease; infertility; vitiligo; drug screening; gene therapy; ss.
 XX XX Homo sapiens.

```

FH Key      Location/Qualifiers
FT CDS      1..2757
FT          /*tag= a
FT          /product= "Human PRIS-18 protein"
FT          /tag= b
FT          /tag= c
FT          /product= "Mature PRIS 18 protein"
FH W0200209496-A2
FH 41-JAN-2002
FH 17-MAY-2002: 2001W0-0522347
FH 21-JUL-2000: 2000US-220064P
FH 28-JUL-2000: 2000US-221680P
FH 04-AUG-2000: 2000US-223544P
FH 11-AUG-2000: 2000US-224717P
FH 16-AUG-2000: 2000US-225986P
FH 23-AUG-2000: 2000US-227568P
FH XX
FH PA        (INCYTE GENOMICS INC)
FH PI        Delegrave AM, Gandhi AR, Hatalia AIA, Lu DAM, Patterson C
FH PI        Tribouley CM, Das D, Kallick DA, Nguyen DB, Lee EA, Khan PA,
FH PI        Yue H, Au-Yang J, Griffin JA, Policky JL, Ramkumar J, Yang J;
FH PI        Thangavelu K, Ding L, Kearney L, Baughn MR, Borowsky ML;
FH PI        Sanjanwalla MS, Yao MG, Burford N, Walla NK, Lal P, Lee S, Todd S;
FH PI        Lo TP, Tang YT, Elliott VS, Azimzai Y, Lu Y;
FH XX
FH DR        WPI: 2002-206083/26.
FH DR        P-PSDB: AAE19181.
FH XX
FH PT        New human protease polypeptide, useful in diagnosis, prevention and
FH PT        treatment of gastrointestinal, cardiovascular, autoimmune/inflammatory,
FH PT        cell proliferative, developmental, epithelial and neurological
FH PT        disorders
FH XX
FH PS        Claim 5: Page 179; 182pp: English.
FH XX
FH CC        The invention relates to an isolated human protease polypeptide (PRIS).
FH CC        PRIS protein and DNA are useful for diagnosis, treating and preventing
FH CC        gastrointestinal disorders (gastritis, cirrhosis, Crohn's disease),
FH CC        autoimmune/inflammatory disorders (AIDS, allergy, rheumatoid arthritis,
FH CC        anemias, asthma), cardiovascular disorder (atherosclerosis, hypertension,
FH CC        myocardial infarction), cell proliferative disorders (hepatitis, cancer,
FH CC        psoriasis), developmental disorders (Cushing's syndrome, hypothyroidism),
FH CC        epithelial disorder (vitiligo, keloid, eczema), neurological disorders
FH CC        (epilepsy, Alzheimer's disease, Pick's disease, Huntington's disease,
FH CC        Parkinson's disease), and reproductive disorders (infertility). PRIS
FH CC        protein is useful in a number of drug screening techniques and to
FH CC        analyse the proteome of a tissue or cell type. PRIS DNA is useful for
FH CC        creating knock-in humanised animals or transgenic animals to model human
FH CC        diseases, in somatic or germline gene therapy and in microarrays
FH CC        utilising fluids or tissues from patients to detect altered PKIN
FH CC        expression. The present sequence is human PRIS-18 cDNA.
FH XX
FH SO        Sequence 2757 BP; 641 A; 802 C; 787 G; 527 T; 0 other.
FH XX

Alignment Scores:
FH Pred. No.:      67.9
FH Score:          9.00
FH Percent Similarity: 100.00%
FH Best Local Similarity: 100.00%
FH Query Match:    2.77%
FH DB:             24
FH Gaps:           0
FH
FH US-09-831-458A-12 (1-325) x AAD30585 (1-2757)
FH
FH UY      34 AlaGlyValLeuValAlaIleLeuVal 42
FH          ||||||||||||||||||||||||||||

```

```

FH DB      2113 GCTGGCATTTGTCGTCGCACTTCATG 2149
FH RESULT 38
FH ID      AAD39159
FH ID      AAD39159 standard; cDNA: 2781 BP.
FH XX
FH AC      AAD39159;
FH XX
FH DT      04-OCT-2002 (first entry)
FH XX
FH CE      Human novel protease cDNA #1.
FH XX
FH KW      Human; novel human protein; NHP; protease; meltrin-beta;
FH KW      metalloprotease; drug screening; clinical trial monitor;
FH KW      pharmacological; nutrient food; gene therapy; chromosome-
FH XX
FH OS      Homo sapiens.
FH XX
FH FH      Key      Location/Qualifiers
FH FT CDS      1..2781
FH FT          /*tag= a
FH FT          /product= "Human novel protease #1"
FH XX
FH PN        W0200236754-A2.
FH XX
FH PD        10-MAY-2002.
FH XX
FH PF        40-OCT-2001: 2001W0-0548448.
FH XX
FH PP        01-NOV-2000: 2000US-244944P.
FH XX
FH PA        (LEXI- ) LEXICON GENETICS INC.
FH PI        Walke DW, Wilkanowski NL, Friedle CJ;
FH XX
FH DR        WPI: 2002-479761/51.
FH DR        P-PSDB: AAE24350.
FH XX
FH PT        New nucleic acid encoding a protein that shares sequence
FH PT        with mammalian meltrin-beta/ADAM 19 homolog metalloprote-
FH PT        diagnosis and treatment, and in drug screening
FH XX
FH PS        Disclosure, Page 34-35; 44pp: English.
FH XX
FH CC        The invention relates to novel human proteins (NHP), novel
FH CC        proteases that shares sequence similarity with mammalian
FH CC        ADAM 19 homologue metalloprotease, and their correspond-
FH CC        sequences. NHP DNA is useful for diagnosis, drug screen-
FH CC        trial monitoring, treatment of diseases and disorders in
FH CC        pharmacological, cosmetic and nutraceutical applications
FH CC        useful as hybridisation probes for screening libraries
FH CC        gene expression patterns (particularly using a microarray
FH CC        throughput chip format). It is also useful in identifying
FH CC        length polymorphism (RFLP) analysis to identify specific
FH CC        and as a probe to identify the corresponding mutant alle-
FH CC        libraries. It is also useful for the detection of mutat-
FH CC        NHPs or inappropriately expressed NHPs for the diagnosis
FH CC        it is also useful in gene therapy. The present sequence
FH CC        protease cDNA. Human NHP gene is located at chromosome-
FH XX
FH SO        Sequence 2781 BP; 644 A; 807 C; 794 G; 534 T; 0 other.
FH XX

Alignment Scores:
FH Pred. No.:      68.5
FH Score:          9.00
FH Percent Similarity: 100.00%
FH Best Local Similarity: 100.00%
FH Query Match:    2.77%
FH DB:             24
FH Gaps:           0
FH
FH US-09-831-458A-12 (1-325) x AAD39159 (1-2781)
FH
FH UY      34 AlaGlyValLeuValAlaIleLeuVal 42
FH          ||||||||||||||||||||||||||||

```

```

|||||
Db 2113 GCTGAGTGTGTGTGGCAGATCTTG 2139
RESULT 39
A039162
1) AAD39162 standard; cDNA: 2868 BP.
XX
AC AAD39162:
XX
DE 04-OCT-2002 (first entry)
XX
DE human novel protease cDNA #4.
XX
KM Human, novel human protein; NHP; protease; meltrin-beta; ADAM 19;
KV metalloprotease; drug screening, clinical trial monitoring, cosmetic,
KW pharmacological, nutraceutical, gene therapy; chromosome 5, gene, SS.
XX
OS Homo sapiens.
XX
FH Key 1..2868 Location/Qualifiers
FT CDS /tag= a
FT /product= "human novel protease #4"
FT
XX
PE W0200236759-A2.
XX
PI 10-MAY-2002.
XX
PE 10-OCT-2001: 2001W0-US48498.
XX
PE 01-NOV-2000: 2000US-244939P.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Walke DW, Wilganowski NL, Friddle CJ.
XX
DE WPI: 2002-479761/51.
XX
DE P-PSDB: AAE24353.
XX
PI New nucleic acid encoding a protein that shares sequence similarity
PI with mammalian meltrin-beta/ADAM 19 homolog metalloprotease for use in
PI diagnosis and treatment, and in drug screening -
XX
PS Disclosure: Page 44; 48pp; English.
XX
CC The invention relates to novel human proteins (NHP), novel human
CC proteases that shares sequence similarity with mammalian meltrin-beta/
CC ADAM 19 homologue metalloprotease, and their corresponding nucleic acid
CC sequences. NHP DNA is useful for diagnosis, drug screening, clinical
CC trial monitoring, treatment of diseases and disorders and
CC pharmacological, cosmetic and nutraceutical applications. It is also
CC useful as hybridisation probes for screening libraries, and assessing
CC gene expression patterns (particularly using a microarray or high-
CC throughput chip format). It is also useful in restriction fragment
CC length polymorphism (RFLP) analysis to identify specific individuals,
CC and as a probe to identify the corresponding mutant allele in the
CC libraries. It is also useful for the detection of mutant
CC NHPs or inappropriately expressed NHPs for the diagnosis of disease.
CC It is also useful in gene therapy. The present sequence is novel human
CC protease cDNA. Human NHP gene is located at chromosome 5.
XX
XX Sequence 2868 BP: 668 A; 826 C; 821 G; 550 T; 0 other:

```

Alignment Scores:

Prod. No.: 70.5 Length: 2868
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.77% Indels: 0
 DB: 24 Gaps: 0

US-09-831-458A-12 (1-325) x AAD39162 (1-2868)

```

OY 34 AAGCAGVallieValAlaIleVal 42
|||||
Db 2113 GCTGAGTGTGTGTGGCAGATCTTG 2139
RESULT 40
A039161
ID AAD39161 standard; cDNA: 2892 BP.
XX
AC AAD39161:
XX
DE 04-OCT-2002 (first entry)
XX
DE human novel protease cDNA #3.
XX
KM Human, novel human protein; NHP; protease; meltrin-beta; ADAM 19;
KV metalloprotease; drug screening, clinical trial monitoring, cosmetic,
KW pharmacological, nutraceutical, gene therapy; chromosome 5, gene, SS.
XX
OS Homo sapiens.
XX
FH Key 1..2892 Location/Qualifiers
FT CDS /tag= a
FT /product= "human novel protease #3"
FT
XX
PE W0200236759-A2.
XX
PI 10-MAY-2002.
XX
PE 10-OCT-2001: 2001W0-US48498.
XX
PE 01-NOV-2000: 2000US-244939P.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Walke DW, Wilganowski NL, Friddle CJ.
XX
DE WPI: 2002-479761/51.
XX
DE P-PSDB: AAE24352.
XX
PI New nucleic acid encoding a protein that shares sequence similarity
PI with mammalian meltrin-beta/ADAM 19 homolog metalloprotease for use in
PI diagnosis and treatment, and in drug screening -
XX
PS Claim 2: Page 40-41; 48pp; English.
XX
CC The invention relates to novel human proteins (NHP), novel human
CC proteases that shares sequence similarity with mammalian meltrin-beta/
CC ADAM 19 homologue metalloprotease, and their corresponding nucleic acid
CC sequences. NHP DNA is useful for diagnosis, drug screening, clinical
CC trial monitoring, treatment of diseases and disorders and
CC pharmacological, cosmetic and nutraceutical applications. It is also
CC useful as hybridisation probes for screening libraries, and assessing
CC gene expression patterns (particularly using a microarray or high-
CC throughput chip format). It is also useful in restriction fragment
CC length polymorphism (RFLP) analysis to identify specific individuals,
CC and as a probe to identify the corresponding mutant allele in the
CC libraries. It is also useful for the detection of mutant
CC NHPs or inappropriately expressed NHPs for the diagnosis of disease.
CC It is also useful in gene therapy. The present sequence is novel human
CC protease cDNA. Human NHP gene is located at chromosome 5.
XX
XX Sequence 2892 BP: 671 A; 834 C; 828 G; 559 T; 0 other:

```

Alignment Scores:

Prod. No.: 71.1 Length: 2892
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.77% Indels: 0
 DB: 24 Gaps: 0

US-09-831-458A-12 (1-325) x AAD39161 (1-2892)

DE Human immune/haematopoietic antigen-presenting sequence. SEQ ID NO: 23250
 XX Human: immune/ haematopoietic immune/haematopoietic antigen-presenting
 KW Cytostatic/ gene therapy/ vaccine/ metastasis/ ds.
 XX Homo sapiens.
 FN W020157182-A2.
 XX
 XX 09-AUG-2001
 PD
 XX 17-JAN-2001: 2001W07501354
 XX
 PR 31-JAN-2000: 2000US-0179665
 PR 04-FEB-2000: 2000US-0180628
 PR 24-FEB-2000: 2000US-0184664
 PR 02-MAR-2000: 2000US-0186360
 PR 16-MAR-2000: 2000US-0189874
 PR 17-MAR-2000: 2000US-0190076
 PR 18-APR-2000: 2000US-0198123
 PR 19-MAY-2000: 2000US-0205515
 PR 07-JUN-2000: 2000US-0209497
 PR 28-JUN-2000: 2000US-0214886
 PR 30-JUN-2000: 2000US-0215135
 PR 07-JUL-2000: 2000US-0216647
 PR 07-JUL-2000: 2000US-0216680
 PR 11-JUL-2000: 2000US-0217487
 PR 11-JUL-2000: 2000US-0217496
 PR 14-JUL-2000: 2000US-0218290
 PR 26-JUL-2000: 2000US-0220963
 PR 26-JUL-2000: 2000US-0220964
 PR 14-AUG-2000: 2000US-0224518
 PR 14-AUG-2000: 2000US-0224519
 PR 14-AUG-2000: 2000US-0225213
 PR 14-AUG-2000: 2000US-0225214
 PR 14-AUG-2000: 2000US-0225267
 PR 14-AUG-2000: 2000US-0225268
 PR 14-AUG-2000: 2000US-0225270
 PR 14-AUG-2000: 2000US-0225447
 PR 14-AUG-2000: 2000US-0225757
 PR 14-AUG-2000: 2000US-0225758
 PR 14-AUG-2000: 2000US-0225759
 PR 18-AUG-2000: 2000US-0226279
 PR 22-AUG-2000: 2000US-0226681
 PR 22-AUG-2000: 2000US-0226686
 PR 22-AUG-2000: 2000US-0227182
 PR 23-AUG-2000: 2000US-0227609
 PR 30-AUG-2000: 2000US-0228624
 PR 01-SEP-2000: 2000US-0229287
 PR 01-SEP-2000: 2000US-0229343
 PR 01-SEP-2000: 2000US-0229344
 PR 01-SEP-2000: 2000US-0229345
 PR 05-SEP-2000: 2000US-0229606
 PR 05-SEP-2000: 2000US-0229613
 PR 06-SEP-2000: 2000US-0230437
 PR 06-SEP-2000: 2000US-0230438
 PR 08-SEP-2000: 2000US-0231242
 PR 08-SEP-2000: 2000US-0231243
 PR 08-SEP-2000: 2000US-0231244
 PR 08-SEP-2000: 2000US-0231413
 PR 08-SEP-2000: 2000US-0231414
 PR 08-SEP-2000: 2000US-0231415
 PR 08-SEP-2000: 2000US-0232081
 PR 12-SEP-2000: 2000US-0231968
 PR 14-SEP-2000: 2000US-0232397
 PR 14-SEP-2000: 2000US-0232398
 PR 14-SEP-2000: 2000US-0232399
 PR 14-SEP-2000: 2000US-0232401
 PR 14-SEP-2000: 2000US-0232401
 PR 14-SEP-2000: 2000US-0233064
 PR 14-SEP-2000: 2000US-0233064
 PR 14-SEP-2000: 2000US-0233065

PR 21-SEP-2000: 2000US-0234228
 PR 21-SEP-2000: 2000US-0234274
 PR 25-SEP-2000: 2000US-0234967
 PR 25-SEP-2000: 2000US-0234998
 PR 26-SEP-2000: 2000US-0235484
 PR 27-SEP-2000: 2000US-0235844
 PR 27-SEP-2000: 2000US-0235846
 PR 29-SEP-2000: 2000US-0236327
 PR 29-SEP-2000: 2000US-0236367
 PR 29-SEP-2000: 2000US-0236368
 PR 29-SEP-2000: 2000US-0236369
 PR 29-SEP-2000: 2000US-0236370
 PR 02-OCT-2000: 2000US-0236802
 PR 02-OCT-2000: 2000US-0237047
 PR 02-OCT-2000: 2000US-0237048
 PR 02-OCT-2000: 2000US-0237049
 PR 02-OCT-2000: 2000US-0237049
 PR 13-OCT-2000: 2000US-0239945
 PR 13-OCT-2000: 2000US-0239947
 PR 29-OCT-2000: 2000US-0240960
 PR 29-OCT-2000: 2000US-0241241
 PR 29-OCT-2000: 2000US-0241785
 PR 29-OCT-2000: 2000US-0241786
 PR 29-OCT-2000: 2000US-0241787
 PR 29-OCT-2000: 2000US-0241808
 PR 29-OCT-2000: 2000US-0241809
 PR 29-OCT-2000: 2000US-0241826
 PR 08-NOV-2000: 2000US-0246474
 PR 08-NOV-2000: 2000US-0246475
 PR 08-NOV-2000: 2000US-0246476
 PR 08-NOV-2000: 2000US-0246477
 PR 08-NOV-2000: 2000US-0246478
 PR 08-NOV-2000: 2000US-0246523
 PR 08-NOV-2000: 2000US-0246524
 PR 08-NOV-2000: 2000US-0246525
 PR 08-NOV-2000: 2000US-0246526
 PR 08-NOV-2000: 2000US-0246527
 PR 08-NOV-2000: 2000US-0246528
 PR 08-NOV-2000: 2000US-0246532
 PR 08-NOV-2000: 2000US-0246609
 PR 08-NOV-2000: 2000US-0246610
 PR 08-NOV-2000: 2000US-0246611
 PR 08-NOV-2000: 2000US-0246613
 PR 17-NOV-2000: 2000US-0249207
 PR 17-NOV-2000: 2000US-0249208
 PR 17-NOV-2000: 2000US-0249209
 PR 17-NOV-2000: 2000US-0249210
 PR 17-NOV-2000: 2000US-0249211
 PR 17-NOV-2000: 2000US-0249212
 PR 17-NOV-2000: 2000US-0249213
 PR 17-NOV-2000: 2000US-0249214
 PR 17-NOV-2000: 2000US-0249215
 PR 17-NOV-2000: 2000US-0249216
 PR 17-NOV-2000: 2000US-0249217
 PR 17-NOV-2000: 2000US-0249218
 PR 17-NOV-2000: 2000US-0249244
 PR 17-NOV-2000: 2000US-0249245
 PR 17-NOV-2000: 2000US-0249246
 PR 17-NOV-2000: 2000US-0249247
 PR 17-NOV-2000: 2000US-0249248
 PR 17-NOV-2000: 2000US-0249249
 PR 01-DEC-2000: 2000US-0250160
 PR 01-DEC-2000: 2000US-0250161
 PR 05-DEC-2000: 2000US-0251040
 PR 05-DEC-2000: 2000US-0251048
 PR 05-DEC-2000: 2000US-0251049
 PR 05-DEC-2000: 2000US-0251179
 PR 08-DEC-2000: 2000US-0251856
 PR 08-DEC-2000: 2000US-0251868
 PR 08-DEC-2000: 2000US-0251869
 PR 08-DEC-2000: 2000US-0251869
 PR 08-DEC-2000: 2000US-0251869

08-DEC-2000: 2000US-0251946
11-DEC-2000: 2000US-0254947
05-JAN-2001: 2001US-0254974
(HOMA-) HUMAN GENOME SCI INC.
Kosen CA, Barash SC, Ruben SM,
WPI: 2001-48426/52
Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
useful for preventing, diagnosing and/or treating cancers and
metastasis -
Disclosure: SEQ ID NO 24260: 3671pp + Sequence Listing: English.
AAK64951 to AAK64702 encode the human immune/hematopoietic antigen (1)
amino acid sequences given in AAK62176 to AAK61941 (1) have cytotoxic
activity, and can be used in gene therapy and vaccine production. (1)
proteins and polynucleotides may be used in the prevention, diagnosis and
treatment of diseases associated with inappropriate (1) expression, but
example, they may be used to treat disorders associated with decreased
expression by rectifying mutations or deletions in a patient's genome
that affect the activity of (1) by expressing inactive proteins or to
supplement the patient's own production of (1). Additionally, (1)
polynucleotides may be used to produce the secreted (1), by inserting
the nucleic acids into a host cell and culturing the cell to express the
protein. (1) proteins and polynucleotides may be used to prevent,
diagnose and treat immune/hematopoietic-related diseases, especially
cancers and cancer metastases of hematopoietic-derived cells. AAK64703
to AAK64704 represent human immune/hematopoietic antigen genomic
sequences from the present invention. AAK54942 to AAK54950 and AAK62159
represent sequences used in the exemplification of the present invention.
Sequence 16424 BP: 4245 A: 3674 C: 4136 G: 4434 T: 0 other.
Alignment Scores:
Prod. No.: 366 Length: 16424
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.77% Indels: 0
DB: 22 Gaps: 0
US-09-831-458a-12 (1) x AAK58448 (1) 2424
CY 34 AAGAGVAILGQVAILALILEVAL 42
|||||
DB 12545 CTTGACCTCTGCTGCTCATCTTCTCTG 12571
RESULT 44
AAK75677/-
ID AAK75677 standard; DNA: 1424 BP.
XX
AC AAK75677:
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/hematopoietic antigen genomic sequence SEQ ID NO 2449.
XX
FM Human immune/hematopoietic immune/hematopoietic antigen; cancer;
KW Cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN W0200157182-A2
XX
PD 09-AUG-2001.
XX
PF 17 JAN 2001: 2001US-0251944
XX
PF 31-JAN-2000: 2000US-0179065
PF 04-FEB-2000: 2000US-0180728

24-MAR-2000: 2000US-0184664.
PR 02-MAR-2000: 2000US-0186350.
PR 16-MAR-2000: 2000US-0186874.
PR 18-MAR-2000: 2000US-0190076.
PR 19-MAY-2000: 2000US-0205515.
PR 07-JUN-2000: 2000US-0205477.
PR 28-JUN-2000: 2000US-0214886.
PR 30-JUN-2000: 2000US-0215135.
PR 07-JUL-2000: 2000US-0216647.
PR 07-JUL-2000: 2000US-0216880.
PR 11-JUL-2000: 2000US-0217487.
PR 11-JUL-2000: 2000US-0217496.
PR 14-JUL-2000: 2000US-0218290.
PR 26-JUL-2000: 2000US-0220964.
PR 26-JUL-2000: 2000US-0220964.
PR 14-AUG-2000: 2000US-0224518.
PR 14-AUG-2000: 2000US-0224519.
PR 14-AUG-2000: 2000US-0225214.
PR 14-AUG-2000: 2000US-0225214.
PR 14-AUG-2000: 2000US-0225266.
PR 14-AUG-2000: 2000US-0225267.
PR 14-AUG-2000: 2000US-0225268.
PR 14-AUG-2000: 2000US-0225270.
PR 14-AUG-2000: 2000US-0225447.
PR 14-AUG-2000: 2000US-0225757.
PR 14-AUG-2000: 2000US-0225758.
PR 14-AUG-2000: 2000US-0225759.
PR 18-AUG-2000: 2000US-0226274.
PR 22-AUG-2000: 2000US-0226281.
PR 22-AUG-2000: 2000US-0226856.
PR 22-AUG-2000: 2000US-0227182.
PR 24-AUG-2000: 2000US-0227094.
PR 30-AUG-2000: 2000US-0228924.
PR 01-SEP-2000: 2000US-0229287.
PR 01-SEP-2000: 2000US-0229344.
PR 01-SEP-2000: 2000US-0229344.
PR 01-SEP-2000: 2000US-0229345.
PR 05-SEP-2000: 2000US-0229509.
PR 05-SEP-2000: 2000US-0229511.
PR 06-SEP-2000: 2000US-0230437.
PR 06-SEP-2000: 2000US-0230438.
PR 08-SEP-2000: 2000US-0231242.
PR 08-SEP-2000: 2000US-0231243.
PR 08-SEP-2000: 2000US-0231244.
PR 08-SEP-2000: 2000US-0231414.
PR 08-SEP-2000: 2000US-0231414.
PR 08-SEP-2000: 2000US-0232080.
PR 08-SEP-2000: 2000US-0232081.
PR 12-SEP-2000: 2000US-0231968.
PR 14-SEP-2000: 2000US-0232397.
PR 14-SEP-2000: 2000US-0232398.
PR 14-SEP-2000: 2000US-0232399.
PR 14-SEP-2000: 2000US-0232400.
PR 14-SEP-2000: 2000US-0232401.
PR 14-SEP-2000: 2000US-0232401.
PR 14-SEP-2000: 2000US-0232404.
PR 14-SEP-2000: 2000US-0233064.
PR 14-SEP-2000: 2000US-0233065.
PR 21-SEP-2000: 2000US-0234224.
PR 21-SEP-2000: 2000US-0234224.
PR 25-SEP-2000: 2000US-0244997.
PR 25-SEP-2000: 2000US-0244998.
PR 26-SEP-2000: 2000US-0244998.
PR 26-SEP-2000: 2000US-0235484.
PR 27-SEP-2000: 2000US-0235834.
PR 27-SEP-2000: 2000US-0235836.
PR 29-SEP-2000: 2000US-0236327.
PR 29-SEP-2000: 2000US-0236327.
PR 29-SEP-2000: 2000US-0236328.
PR 29-SEP-2000: 2000US-0236369.
PR 29-SEP-2000: 2000US-0236370.
PR 02-OCT-2000: 2000US-0236802.
PR 02-OCT-2000: 2000US-0237037.
PR 02-OCT-2000: 2000US-0237038.

PI proteins of insulin receptor substrate family, and signal transducer
 PI and activator of transcription factors to their receptors, useful to
 treat diabetes

PS Disclosure: Page 99-133; 139pp; English.

XX
 CC The invention relates to an isolated pleckstrin homology domain
 CC interacting protein (PHIP) that recruits proteins of the insulin
 CC receptor substrate (IRS) family, and signal transducer and activator of
 CC transcription (STAT) transcription factors, to receptors that interact
 CC with and phosphorylate the proteins and STAT transcription factors,
 CC the nucleic acid encoding PHIP (nPHIP), a nucleic acid which binds to
 CC nPHIP or regions of it, analogues, fragments or allelic variants of PHIP
 CC or nPHIP, a nucleic acid sequence having substantial sequence identity or
 CC sequence similarity with a nucleic acid sequence fully defined
 CC human neuronal differentiation-related protein (NURF) nucleic acid
 CC sequence or its exons as given in the specification, expression
 CC vectors and host cells expressing the nucleic acids, anti-PHP antibodies,
 CC and a transgenic animal not already expressing PHIP. The nucleic
 CC acids, proteins and antibodies are useful for diagnosis and treatment of
 CC a condition associated with an insulin receptor (e.g. diabetes mellitus
 CC type 2, hyperglycaemia, myotonic muscular dystrophy, acanthosis,
 CC nigricans, retinopathy, nephropathy, arteriosclerosis, peripheral
 CC arterial disease) or cancer (e.g. adenocarcinoma, leukaemia, breast
 CC cancer, prostate cancer, colon cancer, ovarian cancer and many others
 CC given in the specification), autoimmune disease, inflammation and
 CC immunodeficiency. The protein is also useful for discovering
 CC or testing compounds which may be either enhancers or inhibitors of PHIP
 CC function. The present sequence is genomic DNA encoding PHIP and NURF.

XX
 SO Sequence 140036 RF, 48506 A, 25843 C, 22459 G, 43134 T, 0 other;

Alignment Scores:

Pred. No.:	2	77e+03	Length:	140036
Score:	9.00		Matches:	9
Percent Similarity:	100.00%		Conservative:	0
Best Local Similarity:	100.00%		Mismatches:	0
Query Match:	2.77%		Indels:	0
DB:	24		Gaps:	0

US-09-831-458a-12 (1-345) X AAS9800U (1-140036)

OY 74 G1ULUUSerG1ULYSerLysLeuGln 82

DB 107457 GAGTAAAGTGAAGAAATCTAACTACAA 107431

Search completed: December 7, 2002, 13:53:17
 Job time : 353 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd

OM protein - nucleic search, using frame-plus-F2u model

Run on: December 7, 2002, 11:31:17 Search time: 55 seconds
(without alignments)

1812.181 Million cell updates/sec

Title: US-09-831-458a-12

Sequence: 325 1 MSLSKPRVQQLGLGCTGH LCTVINWVTKKPAAPM E 325

Scoring table:

OLIGO	60.0	Xgapext 60.0
Xgapop 60.0	Xgapext 60.0	
fgapop 6.0	fgapext 7.0	
delop 6.0	delext 7.0	

Searched: 44162 seqs, 15338381 residues

Word size: 5

Total number of hits satisfying chosen parameters: 32521

Minimum db seq length: 0

Maximum db seq length: 20000000

Post-processing: listing first 45 summaries

Command line parameters:

-MODEL=frame-p2u.model -DEV=xlh
-U/cgm2_1/70SPN.spool/US09831458/runat_05122002_103417_16789/apf_query.fasta_1.519
-DB=Issued_Patents_NA -OFMT=fastap -SUFFIX=oli.rmi -MINMATCH=0 -1 -100pct=0
-LOOPEXT=0 -ONITS-bits -START=1 -END=1 -MATRIX=oligo -TRANS=humand40.cdi
-LIST=45 -LOCALIGN=200 -THR_SCORE=quality -THR_MIN=5 -ALIGN=45 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=40 -MAXLEN=20000000
-USPR=US09831458.acrn_1.1.17 -runat_05122002_103417_16789 -NCPV 6 -TRF=3
-NO_XLPRY -NO_MMAR -LARGEQUERY -NEG_SCORES=0 -WAIT -LONELUG -DEV_TIMECOUT=120
-WARN_TIMECOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YCAPOP=60 -YCAPEXT=60 -DELCP=6 -DELEXT=7

Database: Issued_Patents_NA.*

1: /cgm2_6/prodata/1/ina/5A.COMB.seq.*
2: /cgm2_6/prodata/1/ina/5B.COMB.seq.*
3: /cgm2_6/prodata/1/ina/5A.COMB.seq.*
4: /cgm2_6/prodata/1/ina/5B.COMB.seq.*
5: /cgm2_6/prodata/1/ina/5A.COMB.seq.*
6: /cgm2_6/prodata/1/ina/5B.COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	DB ID	Description
1	168	51.7	1643	US-09-517-605-16
2	78	24.0	1212	US-09-591-435-3
3	78	24.0	1312	US-09-517-605-1
4	43	13.2	4	US-09-591-435-10
5	35	10.8	1212	US-09-591-435-11
6	9	2.8	606	US-09-280-116-11
7	9	2.8	2251	US-08-836-443-2
8	9	2.8	2648	US-08-836-443-1
9	8	2.5	3744	US-08-348-353-16
10	8	2.5	3744	US-08-348-353-16
11	8	2.5	3744	US-08-465-965-16
12	8	2.5	8075	US-08-465-966-16
				US-08-374-077C-1
				Sequence 1, Appl

13	8	2.5	8075	US-08-835-690-1	Appl
14	8	2.5	8075	US-09-539-879A-1	Appl
15	7	2.2	132	US-08-450-945-68	Appl
16	7	2.2	132	US-08-476-161-68	Appl
17	7	2.2	360	US-07-744-518-24	Appl
18	7	2.2	407	US-09-280-116-105	Appl
19	7	2.2	452	US-07-662-198A-1	Appl
20	7	2.2	452	US-08-342-142-1	Appl
21	7	2.2	462	US-08-450-945-65	Appl
22	7	2.2	462	US-08-476-161-66	Appl
23	7	2.2	469	US-08-458-084-5	Appl
24	7	2.2	469	US-08-205-508-5	Appl
25	7	2.2	469	US-08-182-148-2	Appl
26	7	2.2	469	US-08-459-967-9	Appl
27	7	2.2	469	US-08-459-967-9	Appl
28	7	2.2	538	US-08-031-148-1	Appl
29	7	2.2	538	US-08-415-838-1	Appl
30	7	2.2	538	US-07-205-169-1	Appl
31	7	2.2	612	US-09-385-982-64	Appl
32	7	2.2	679	US-08-134-255-9	Appl
33	7	2.2	679	US-08-459-967-9	Appl
34	7	2.2	679	US-08-460-327-9	Appl
35	7	2.2	679	US-08-459-967-9	Appl
36	7	2.2	690	US-09-419-548E-24	Appl
37	7	2.2	690	US-09-454-248E-24	Appl
38	7	2.2	763	US-08-468-084-4	Appl
39	7	2.2	763	US-08-205-508-4	Appl
40	7	2.2	763	US-08-482-148-1	Appl
41	7	2.2	763	US-08-482-148-1	Appl
42	7	2.2	763	US-08-482-148-1	Appl
43	7	2.2	791	US-09-051-566-5	Appl
44	7	2.2	936	US-09-416-509C-2	Appl
45	7	2.2	982	US-08-973-068-28	Appl

ALIGNMENTS

RESULT 1
US-09-517-605-16
Sequence 16, Application: US-09-517-605-16
Patent No. 6391567
GENERAL INFORMATION:
APPLICANT: Liltman, Dan R.
APPLICANT: Kwon, Douglas S.
APPLICANT: van Kooyk, Yvette
TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RE:
FILE REFERENCE: 1049-1-017
CURRENT APPLICATION NUMBER: US-09-517-605-16
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.0
Seq ID No 16
LENGTH: 1643
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: CDS
LOCATION: (25)..
US-09-517-605-16
Alignment Scores:
Pred. No.: 3.240E+00
Score: 168.00
Percent Similarity: 100.00
Best Local Similarity: 100.00
Query Match: 51.698
DB: 4
US-09-831-458A-12 (1-325) X US-09-517-605-16 (1-1643)
1 MetSerAspSerIlySGIupProAlaValGlnIleuGlyIleuLeuGly

```

1b 25 ATGAGTGAATCTGAAAGAAATCAAGGCTGAGTGGGCTTCTGAGAGGCTTGGAGAT 64
QY 21 CTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 46
1b 85 GGGGCCCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 144
QY 41 TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
1b 145 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 204
QY 61 TTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 80
1b 205 ATCTACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 264
QY 81 TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 100
1b 265 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 324
QY 101 TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
1b 325 AAATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 384
QY 121 TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 140
1b 385 TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 444
QY 141 TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 160
1b 445 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 504
QY 161 TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 168
1b 505 AAATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 528

```

```

RESULT 2
US-09-591-435-9
: Sequence 9, Application US/09591435
: Patent No. 6280953
: GENERAL INFORMATION:
: APPLICANT: MESSIER, WALTER
: APPLICANT: SIKELA, JAMES M
: TITLE OF INVENTION: METHODS TO IDENTIFY POLYNUCLEOTIDE AND POLYPEPTIDE
: TITLE OF INVENTION: SEQUENCES WHICH MAY BE ASSOCIATED WITH PHYSIOLOGICAL
: FILE REFERENCE: GENO 200 2
: CURRENT APPLICATION NUMBER: US/09/591.435
: PRIOR FILING DATE: 2000-06-09
: PRIOR APPLICATION NUMBER: 09/591.435
: PRIOR FILING DATE: 2000-06-09
: PRIOR APPLICATION NUMBER: 09/240.915
: PRIOR FILING DATE: 1999-01-29
: PRIOR APPLICATION NUMBER: 60/073.263
: PRIOR FILING DATE: 1998-01-30
: PRIOR APPLICATION NUMBER: 60/098.987
: PRIOR FILING DATE: 1998-04-02
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 9
: LENGTH: 1212
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-591-435-9

```

```

Alignment Scores:
Pred. No. 1 38e-64 Length: 1212
Score: 78.00 Matches: 78
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 24.00% Indels: 0
DB: 4 Gaps: 0

```

US-09-831-458a-12 (1-325) x US-09-591-435-9 (1 1212)

```

QY 58 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 77
1b 226 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 279
QY 78 TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 97
1b 280 AAATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 339
QY 98 TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 117
1b 440 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 399
QY 118 TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 135
1b 440 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 453

```

```

RESULT 3
US-09-517-605-1
: Sequence 1, Application US/09517605
: Patent No. 6391567
: GENERAL INFORMATION:
: APPLICANT: Littman, Dan R.
: APPLICANT: Kwan, Douglas S.
: APPLICANT: van Koyk, Yvette
: APPLICANT: Gelfondbeck, Theo
: TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
: TITLE OF INVENTION: CELLS
: FILE REFERENCE: 1049-1-017
: CURRENT APPLICATION NUMBER: US/09/517.605
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 1312
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (42)..(1253)
US-09-517-605-1

```

```

Alignment Scores:
Pred. No. 1 48e-64 Length: 1312
Score: 78.00 Matches: 78
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 24.00% Indels: 0
DB: 4 Gaps: 0

```

US-09-831-458a-12 (1-325) x US 09 517-605-1 (1-1312)

```

QY 58 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 77
1b 261 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 329
QY 78 TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 97
1b 321 AAATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 380
QY 98 TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 117
1b 381 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 440

```

```

QY 118 TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 135
1b 441 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 494

```

```

RESULT 4
US-09-591-435-10
: Sequence 10, Application US/09591435
: Patent No. 6280953
: GENERAL INFORMATION:

```



```

COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,443
FILING DATE: 01-MAY-1997
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 9709420
FILING DATE: 05-SEP-1997
APPLICATION NUMBER: 9612145.4
FILING DATE: 11-JUN-1996
APPLICATION NUMBER: 9526230.9
FILING DATE: 21-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 47424
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2251 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-836-443-2

Alignment Scores:
Pred. No. 9 29
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.77%
Indels: 0
Gaps: 0

US-09-831-458a-12 (1-325) x US-08-836-443 (1-2251)
QY 34 AlaclyValleuValaAlaAlleuVal 42
|||||
DB 1586 GCGGCAATCTTCTGCGCATTTCTCG 1613

RESULT 8
US-08-836-443-1
Sequence 1, Application US/08836443
Patent No. 5883241
GENERAL INFORMATION:
APPLICANT: DOCHERTY, Andrew, J.P
APPLICANT: SLOCOMBE, Patrick, M.
TITLE OF INVENTION: DNA SEQUENCES CODING FOR A
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,443
FILING DATE: 01-MAY-1997

```

```

CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 9709420
FILING DATE: 05-SEP-1997
APPLICATION NUMBER: 9612145.4
FILING DATE: 11-JUN-1996
APPLICATION NUMBER: 9526230.9
FILING DATE: 21-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 47424
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2648 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-836-443-1

Alignment Scores:
Pred. No. 10 8
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.77%
Indels: 0
Gaps: 0

US-09-831-458a-12 (1-325) x US-08-836-443 (1-2648)
QY 34 AlaclyValleuValaAlaAlleuVal 42
|||||
DB 1590 GCGGCAATCTTCTGCGCATTTCTCG 1616

RESULT 9
US-08-836-443-16/6
Sequence 16, Application US/08348353
Patent No. 5932217
GENERAL INFORMATION:
APPLICANT: Thomanen, Elaine
APPLICANT: Masure, Robert
TITLE OF INVENTION: Antibody Recognizing Endothelial Cell
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/348,353
FILING DATE: 30-NOV-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Jackson, David
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-097C1P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-1684
TELEFAX: 201-443-1684
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:

```


LENGTH: 3744 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3744
PUBLICATION INFORMATION:
AUTHORS: Delisse-Gathoye, et al.
TITLE: Cloning, Partial Sequence, Expressions, and
TITLE: Antigenic Analysis of the Filamentous
TITLE: Hemagglutinin Gene of Bordetella Pertussis
JOURNAL: Infection and Immunity
VOLUME: 58
ISSUE: 9
PAGES: 2895-2905
DATE: September-1990
US-08-348-453-16

Alignment Scores:
Pred. No.: 152 Length: 3744
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2 46% Indels: 0
DB: 2 Gaps: 0

US-09-831-458a-12 (1-325) x US-08-348-353-16 (1-3744)

QY 36 ValLeuValAlaIleLeuValGln 43
|||||
Db 3291 GTACTCGTAGCGATTCTGTCGACG 3268

RESULT 10
US-08-465-965-16/C
Sequence 16, Application US/08465965
Patent No. 5968512
GENERAL INFORMATION:
APPLICANT: Tuomanen, Elaine
APPLICANT: Masure, Robert
TITLE OF INVENTION: Antibody Recognizing Endothelial Cell
TITLE OF INVENTION: Ligand for Leukocyte CR3
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,965
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/348,353
FILING DATE: 30-NOV-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/247,572
FILING DATE: 23-MAY-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/03725
FILING DATE: 04-MAY-1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/695,613

FILING DATE: 03-MAY-1991
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Jackson, David
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-0971P111V2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 3744 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3744
PUBLICATION INFORMATION:
AUTHORS: Delisse-Gathoye, et al.
TITLE: Cloning, Partial Sequence, Expressions, and
TITLE: Antigenic Analysis of the Filamentous
TITLE: Hemagglutinin Gene of Bordetella Pertussis
JOURNAL: Infection and Immunity
VOLUME: 58
ISSUE: 9
PAGES: 2895-2905
DATE: September-1990
US-08-465-965-16

Alignment Scores:
Pred. No.: 152 Length: 3744
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2 46% Indels: 0
DB: 2 Gaps: 0

US-09-831-458a-12 (1-325) x US-08-465-965-16 (1-3744)

QY 36 ValLeuValAlaIleLeuValGln 43
|||||
Db 3291 GTACTCGTAGCGATTCTGTCGACG 3268

RESULT 11
US-08-465-966-16/C
Sequence 16, Application US/08465966
Patent No. 6015560
GENERAL INFORMATION:
APPLICANT: Tuomanen, Elaine
APPLICANT: Masure, Robert
TITLE OF INVENTION: Antibody recognizing Endothelial
TITLE OF INVENTION: Ligand for Leukocyte CR3
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,966
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/348,354
FILING DATE: 30-NOV-1994

```

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/247,572
  FILING DATE: 23-MAY-1994
  CLASSIFICATION: 424
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: PCT/US92/07725
    FILING DATE: 04-MAY-1992
    CLASSIFICATION: 424
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 07/695,613
    FILING DATE: 03-MAY-1991
    CLASSIFICATION: 424
  ATTORNEY/AGENT INFORMATION:
    NAME: Jackson, David
    REGISTRATION NUMBER: 26,742
    REFERENCE/DOCKET NUMBER: 600-1-00071P1D1V2
    TELEPHONE: 201-487-5800
    TELEFAX: 201-343-1684
  INFORMATION FOR SEQ ID NO: 16:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 3744 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: linear
  FEATURE:
    NAME/KEY: CDS
    LOCATION: 1..3744
  PUBLICATION INFORMATION:
    AUTHORS: Delisse-Gathoye, et al.
    TITLE: Cloning, Partial Sequence, Expressions, and
    TITLE: Antigenic Analysis of the Filamentous
    TITLE: Hemagglutinin Gene of Bordetella Pertussis
    JOURNAL: Infection and Immunity
    VOLUME: 58
    ISSUES: 9
    PAGES: 2895-2905
    DATE: September-1990
  US-08-465-966-16
  Alignment Scores:
    Pred. No.: 152
    Score: 8.00
    Percent Similarity: 100.00%
    Best Local Similarity: 100.00%
    Query Match: 2.46%
    DE: 3
    Caps: 0
  US-09-831-458a-12 (1-325) x US-08-465-966-16 (1-3744)
  QY 36 ValdeuValAlAldeuValgin 43
  DB 3291 GTCCTGTAGCGATCTGTGTCTGAG 3268
  RESULT 12
  US-08-474-077C-1
  Sequence 1, Application US/08374077C
  Patent No. 6027912
  GENERAL INFORMATION:
    APPLICANT: Hall, Linda M
    APPLICANT: Ren, Dejian
    APPLICANT: Zheng, Wei
    APPLICANT: Dupald, Manuel Marcel Paul
    TITLE OF INVENTION: Genes Encoding an Invertebrate Alpha1
    TITLE OF INVENTION: Calcium Channel Subunit
    NUMBER OF SEQUENCES: 57
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: BURNS, DOANE, SWICKER & MATHIS, LLP
    STREET: 699 Prince Street
    CITY: Alexandria
    STATE: VA
    COUNTRY: USA

```

```

  Filing: 22414-3187
  COMPUTER READABLE FORM:
    MEDIA TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.30
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/374,077C
    FILING DATE: 19-JAN-1995
    CLASSIFICATION: 435
  ATTORNEY/AGENT INFORMATION:
    NAME: McCowan, Malcolm M.
    REGISTRATION NUMBER: 39,300
    REFERENCE/DOCKET NUMBER: 022650-264
    TELEPHONE: 703 836 6620
    TELEFAX: 703 836 2921
  INFORMATION FOR SEQ ID NO: 1:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 8075 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: linear
  MOLECULE TYPE: DNA (genomic)
  FEATURE:
    NAME/KEY: CDS
    LOCATION: 157..7704
  US-08-374-077C-1
  Alignment Scores:
    Pred. No.: 313
    Score: 8.00
    Percent Similarity: 100.00%
    Best Local Similarity: 100.00%
    Query Match: 2.46%
    DB: 3
    Caps: 0
  US-09-831-458a-12 (1-325) x US-08-374-077C-1 (1-8075)
  QY 16 GTCCTGTAGCGATCTGTGTCTGAG 23
  DB 3153 GTCCTGTAGCGATCTGTGTCTGAG 3176
  RESULT 13
  US-08-845-590-1
  Sequence 1, Application US/08845590
  Patent No. 6207410
  GENERAL INFORMATION:
    APPLICANT: Hall, Linda M.
    APPLICANT: Ren, Dejian
    APPLICANT: Zheng, Wei
    APPLICANT: Dupald, Manuel Marcel Paul
    TITLE OF INVENTION: Genes encoding an insect Calcium Channel
    NUMBER OF SEQUENCES: 101
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: BURNS, DOANE, SWICKER & MATHIS, LLP
    STREET: 699 Prince Street
    CITY: Alexandria
    STATE: VA
    COUNTRY: USA
    Filing: 22414-3187
  COMPUTER READABLE FORM:
    MEDIA TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.30
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/845,590
    FILING DATE:
    CLASSIFICATION:
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US/08/374,888
      FILING DATE: 19-JAN-1995

```

ATTORNEY/AGENT INFORMATION:
NAME: MCGOWAN, MALCOLM M.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 022650-263
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6620
TELEFAX: 703-836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8075 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 157..7704
US-09-831-458A-12 (1-325) x US-08-895-590-1 (1-8075)
Alignment Scores:
Pred. No.: 313 Length: 8075
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.46% Indels: 0
DB: 4 Gaps: 0
US-09-831-458A-12 (1-325) x US-08-895-590-1 (1-8075)
UY 16 C1GCTCTGCTTGGCGACGACGACATTA 3176
DB 3153 GGGTGTCTTGGCGACGACGACATTA 3176
RESULT 14
US-09-539-879A-1
Sequence 1, Application US/09539879A
Patent No. 6436627
GENERAL INFORMATION:
APPLICANT: Hall, Linda M.
Ken, Dejian
Zheng, Wei
Dubald, Manuel Marcel Paul
TITLE OF INVENTION: Genes Encoding an Invertebrate Alpha-
Calcium Channel Subunit
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: HURNS, DOANE, SWECKER & MATHIS, LLP
STREET: 609 Prince Street
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22314-3187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/539,879A
FILING DATE: 31-Mar-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/111,865
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/374,077
FILING DATE: 19-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: MCGOWAN, MALCOLM M.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 022650-264
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6620
TELEFAX: 703-836-2021

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8075 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 157..7704
US-09-539-879A-1
Alignment Scores:
Pred. No.: 313 Length: 8075
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.46% Indels: 0
DB: 4 Gaps: 0
US-09-831-458A-12 (1-325) x US-09-539-879A-1 (1-8075)
UY 16 C1GCTCTGCTTGGCGACGACGACATTA 3176
DB 3153 GGGTGTCTTGGCGACGACGACATTA 3176
RESULT 15
US-08-450-945-68
Sequence 68, Application US/08450945
Patent No. 5783383
GENERAL INFORMATION:
APPLICANT: Kondo, Kazuhito
Mocarski, Edward S, JR.
TITLE OF INVENTION: LATENT TRANSCRIPTS AND PROMOTERS
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: DeHlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,945
FILING DATE: 23-MAY-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Shultz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 8600-0157
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 132 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: art isolate chr 3
FEATURE:
NAME/KEY: CDS

```

LOCATION: 1.132
US-08-450 945-68

Alignment Scores:
Pred. No.: 68
Score: 7.00
Length: 132
Percent Similarity: 100.00%
Conservative: 7
Best Local Similarity: 100.00%
Mismatch: 0
Query Match: 2.15%
Indels: 0
DB: 1
Gaps: 0

US-09-831-458A-12 (1-325) x US-08-450-945-68 (1-132)
QY 35 G1YVAl1GnGln1pouGlyLeu 41
|||||
DB 42 GGGCTTCCTGCTGCAATCTTC 62

RESULT 16
US-08-976-161-68
Sequence 68, Application US/08976161
Patent No. 6194542
GENERAL INFORMATION:
APPLICANT: Kondo, Kazuhiro
APPLICANT: Mocarisi, Edward S. Jr.
TITLE OF INVENTION: LATENT 1KANSCHIPS AND PROMOTERS
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESS: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/976,161
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/450,945
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/WORKET NUMBER: 8600-0157
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 132 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYBOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: antisense OPG 4
FEATURE:
NAME/KEY: CDS
LOCATION: 1..132
US-08-976-161-68

Alignment Scores:
Pred. No.: 68
Score: 7.00
Length: 132
Percent Similarity: 100.00%
Conservative: 7
Mismatch: 0
Query Match: 0
DB: 0
Gaps: 0

```

```

Best Local Similarity: 100.00%
Query Match: 2.15%
Indels: 0
DB: 4
Gaps: 0

US-09-831-458A-12 (1-325) x US-08-976-161-68 (1-132)
QY 35 G1YVAl1GnGln1pouGlyLeu 41
|||||
DB 42 GGGCTTCCTGCTGCAATCTTC 62

RESULT 17
US-07-743-518-23
Sequence 23, Application US/07743518
Patent No. 5397696
GENERAL INFORMATION:
APPLICANT: YANAGIHARA, RICHARD
APPLICANT: NEROKAR, VIVEK R.
APPLICANT: JENKINS, CAROL
APPLICANT: MILLER, MARK
APPLICANT: GAREPETO, PALPH M.
TITLE OF INVENTION: PAPUA NEW GUINEA HUMAN T-LYMPHOTROPIC
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESS: CUSHMAN, DARVY & CUSHMAN
STREET: 1615 L STREET, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20046-5601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/743,518
FILING DATE: 19910812
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: SCOTT, MATSON T.
REGISTRATION NUMBER: 26,581
REFERENCE/WORKET NUMBER: WTS/5683/84609/SAP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 20..443
US-07-743-518-23

Alignment Scores:
Pred. No.: 178
Score: 7.00
Length: 369
Percent Similarity: 100.00%
Conservative: 7
Best Local Similarity: 100.00%
Mismatch: 0
Query Match: 2.15%
Indels: 0
DB: 1
Gaps: 0

US-09-831-458A-12 (1-325) x US-07-743-518-23 (1-369)
QY 8 ArYVAl1GnGln1pouGlyLeu 14
|||||
DB 60 AGASTAGAGCAG31GGGGCTTC 40

```

RESULT 18
US-09-280-116-105
Sequence 105, Application US/09280116A
Patent No. 6331427
GENERAL INFORMATION:
APPLICANT: Robison, Keith E.
TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
FILE REFERENCE: 5800-24, 035800/176965
CURRENT APPLICATION NUMBER: US/09/280,116A
CURRENT FILING DATE: 1999-03-26
NUMBER OF SEQ ID NOS: 268
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 105
LENGTH: 407
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: trypsin-like serine proteases
NAME/KEY: misc_feature
LOCATION: (1)-(407)
OTHER INFORMATION: n = a, t, c or g
US-09-280-116-105

Alignment Scores:
Pred. No.: 195 Length: 407
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 4 Gaps: 0

US-09-831-458a-12 (1-325) x US-09-280-116-105 (1-407)

OY 10 GlnGlnLeuGlyLeuGly 16
|||||
Db 114 CAGCACTGGGCTTACTAGGG 134

RESULT 19
US-07-662-198B-1/C
Sequence 1, Application US/07662198B
Patent No. 5262528
GENERAL INFORMATION:
APPLICANT: Sager, Ruth
APPLICANT: Lee, Sam W.
APPLICANT: Tomasetto, Catherine
TITLE OF INVENTION: CANCER DIAGNOSIS AND THERAPY
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/662,198B
FILING DATE: 19910228
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 00530/049001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070

TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 452
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
US-07-662-198B-1

Alignment Scores:
Pred. No.: 215 Length: 452
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 1 Gaps: 0

US-09-831-458a-12 (1-325) x US-07-662-198B-1 (1-452)

OY 32 LeuLeuAlaGlyValLeuVal 48
|||||
Db 116 CTCTGGCAGCTACTTGTG 96

RESULT 20
US-08-322-742-1/C
Sequence 1, Application US/08322742
Patent No. 5688641
GENERAL INFORMATION:
APPLICANT: Sager, Ruth
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/322,742
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/948,848
FILING DATE: September 1, 1992
APPLICATION NUMBER: 07/944,296
FILING DATE: February 28, 1992
APPLICATION NUMBER: 07/552,216
FILING DATE: February 28, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 00530/048003
TELECOMMUNICATION INFORMATION:
TELEFAX: (617) 542-8906
TELEPHONE: (617) 542-5070
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 452
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-322-742-1
Alignment Scores:
Pred. No.: 215 Length: 452
Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 1 Gaps: 0

US-09-831-458a-12 (1-325) x US-08-450-945-66 (1-462)

UY 32 LepteqalaglyValleuVal 38
|||||
DB 116 CTTCTGCCAGACTACTCTG 96

RESULT 21

US-08-450-945-66
Sequence 66, Application US/08450945
Patent No. 5783383

GENERAL INFORMATION:

APPLICANT: Kondo, Kazuhiko
APPLICANT: Mocarski, Edward S. Jr.
TITLE OF INVENTION: OF CYTOMEGALOVIRUS
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSER: Deninger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentia Release #1 0, Version #1 30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/450,945

FILING DATE: 23-MAY-1995

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Shultz, Charles K

REGISTRATION NUMBER: 38,615

REFERENCE/DOCKET NUMBER: 8600-0157

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 66:

SEQUENCE CHARACTERISTICS:

LENGTH: 462 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: antisense ORF 2

FEATURE:

NAME/KEY: CDS

LOCATION: 1..462

US-08-450-945-66

Alignment Scores:
Pred. No.: 219 Length: 462
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 1 Gaps: 0

US-09-831-458a-12 (1-325) x US-08-450-945-66 (1-462)

UY 35 CTVallleValAlJleleu 41
|||||
DB 67 GGGTTCTGTCGCAATCTG 87

RESULT 22

US-08-976-161-66

Sequence 66, Application US/08976161

Patent No. 6194542

GENERAL INFORMATION:

APPLICANT: Kondo, Kazuhiko

APPLICANT: Mocarski, Edward S. Jr.

TITLE OF INVENTION: LATENT TRANSCRIPTS AND PROMOTERS

NUMBER OF SEQUENCES: 75

CORRESPONDENCE ADDRESS:

ADDRESSER: Deninger & Associates

STREET: 350 Cambridge Avenue, Suite 250

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentia Release #1 0, Version #1 30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/976,161

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/450,945

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Shultz, Charles K

REGISTRATION NUMBER: 38,615

REFERENCE/DOCKET NUMBER: 8600-0157

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 66:

SEQUENCE CHARACTERISTICS:

LENGTH: 462 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: antisense ORF 2

FEATURE:

NAME/KEY: CDS

LOCATION: 1..462

US-08-976-161-66

Alignment Scores:
Pred. No.: 219 Length: 462
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 4 Gaps: 0

US-09-831-458a-12 (1-325) x US-08-976-161-66 (1-462)

UY 35 CTVallleValAlJleleu 41
|||||
DB 67 GGGTTCTGTCGCAATCTG 87

RESULT 23

US-08-458-084-5

Sequence 5, Application US/08458084

Patent No. 5624837

GENERAL INFORMATION:

APPLICANT: Podor, William L

APPLICANT: Rollins, Stephen P
TITLE OF INVENTION: Chimeric Complement
TITLE OF INVENTION: Inhibitor Proteins
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maurice M. Klee
STREET: 1951 Burr Street
CITY: Fairfield
STATE: Connecticut
COUNTRY: USA
ZIP: 06430
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 750 kb storage
COMPUTER: Dell 486/50
OPERATING SYSTEM: DOS 6.2
SOFTWARE: WordPerfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,084
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Klee, Maurice M.
REGISTRATION NUMBER: 30,399
REFERENCE/DOCKET NUMBER: ALX-120
TELEPHONE: (203) 254-1400
TELEFAX: (203) 254-1101
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Linear
MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: AGCIP full length cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Cercopithecus aethiops
CELL LINE: COS-1 (ATCC CRL 1650)
US-08-458-084-5
Alignment Scores:
Pred. No.: 223 Length: 469
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
Gaps: 0
US-09-831-458A-12 (1-325) x US-08-458-084-5 (1-469)
OY 47 ValProserSerLeuSerGln 53
Db 383 GTGCTTATCCCTAAGTCAA 403
RESULT 24
US-08-205-508-5
Sequence 5, Application US/08205508
Patent No. 5627264
GENERAL INFORMATION:
APPLICANT: Fodor, William L
APPLICANT: Rollins, Stephen P
APPLICANT: Squinto, Stephen P
TITLE OF INVENTION: Chimeric Complement
TITLE OF INVENTION: Inhibitor Proteins
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maurice M. Klee
STREET: 1951 Burr Street
CITY: Fairfield
STATE: Connecticut

COUNTRY: USA
ZIP: 06430
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 750 kb storage
COMPUTER: Dell 486/50
OPERATING SYSTEM: DOS 6.2
SOFTWARE: WordPerfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/205,508
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Klee, Maurice M.
REGISTRATION NUMBER: 30,399
REFERENCE/DOCKET NUMBER: ALX-120
TELEPHONE: (203) 254-1400
TELEFAX: (203) 254-1101
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Linear
MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: AGCIP full length cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Cercopithecus aethiops
CELL LINE: COS-1 (ATCC CRL 1650)
US-08-205-508-5
Alignment Scores:
Pred. No.: 223 Length: 469
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
Gaps: 0
US-09-831-458A-12 (1-325) x US-08-205-508-5 (1-469)
OY 47 ValProserSerLeuSerGln 54
Db 383 GTGCTTATCCCTAAGTCAA 403
RESULT 25
US-08-482-148-2
Sequence 2, Application US/08482148
Patent No. 5847082
GENERAL INFORMATION:
APPLICANT: Rother, Russell
APPLICANT: Rollins, Stephen P
APPLICANT: Squinto, Stephen P
TITLE OF INVENTION: Terminal Complement
TITLE OF INVENTION: Inhibitor Fusion Genes and Proteins
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seth A. Fidler
STREET: Alexion Pharmaceuticals, 45 Science Park
CITY: New Haven
STATE: Connecticut
COUNTRY: USA
ZIP: 06511
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.4 Mb storage
COMPUTER: IBM compatible (Pentium)
OPERATING SYSTEM: Windows 95 under MS D OS
SOFTWARE: Microsoft Word for Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,148
FILING DATE: 6/07/95

CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Fiedel, Seth A.
REGISTRATION NUMBER: 38,449
REFERENCE/DOCKET NUMBER: ALX-129.1 Div
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 776-1790
TELEFAX: (203) 772-3655
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Linear
MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: AGMCIP full length cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Cercarial dermatitis
CELL LINE: COS-1 (ATCC CRL 1650)
US-09-482-148-2

Alignment Scores:
Pred. No.: 223 Length: 469
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: Gaps: 0

US-09-831-458a-12 (1-325) x US-09-482-148-2 (1-469)

Cy 47 ValProSerSerLeuSerGln 53
|||||
Db 383 GTGCTTCATCCTTAGTCA 403

RESULT 26
FCT-US95-02944-2
Sequence 2, Application PC/TMS9502944
GENERAL INFORMATION:
APPLICANT: Rollins, Russell
APPLICANT: Squino, Stephen P
TITLE OF INVENTION: Terminal Complement
TITLE OF INVENTION: Inhibitor Fusion Genes and Proteins
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maurice M. Klee
STREET: 1951 Burr Street
CITY: Fairfield
STATE: Connecticut
COUNTRY: USA
ZIP: 06430
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 720 kb storage
COMPUTER: Dell 486/50
OPERATING SYSTEM: DOS 6.2
SOFTWARE: Word Perfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02944
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/295,720
FILING DATE: 3-MAR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Klee, Maurice M.
REGISTRATION NUMBER: 30,399
REFERENCE/DOCKET NUMBER: ALX-129FCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 255-1400

TELEFAX: (203) 254-1101
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Linear
MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: AGMCIP full length cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Cercarial dermatitis
CELL LINE: COS-1 (ATCC CRL 1650)
FCT US95-02944-2

Alignment Scores:
Pred. No.: 223 Length: 469
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: Gaps: 0

US-09-831-458a-12 (1-325) x PCT US95-02944-2 (1-469)

Cy 47 ValProSerSerLeuSerGln 53
|||||
Db 383 GTGCTTCATCCTTAGTCA 403

RESULT 27
FCT-US95-02945-5
Sequence 5, Application PC/TMS9502945
GENERAL INFORMATION:
APPLICANT: Fodor, William L.
APPLICANT: Rollins, Stephen P
APPLICANT: Squino, Stephen P
TITLE OF INVENTION: ChimERIC Complement
TITLE OF INVENTION: Inhibitor Proteins
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maurice M. Klee
STREET: 1951 Burr Street
CITY: Fairfield
STATE: Connecticut
COUNTRY: USA
ZIP: 06430
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 720 kb storage
COMPUTER: Dell 486/50
OPERATING SYSTEM: DOS 6.2
SOFTWARE: WordPerfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02945
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/295,508
FILING DATE: 3-MAR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Klee, Maurice M.
REGISTRATION NUMBER: 30,399
REFERENCE/DOCKET NUMBER: ALX-120PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 255-1400
TELEFAX: (203) 254-1101
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Linear

MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: AGC1P full length cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Cercopithecus aethiops
CELL LINE: COS-1 (ATCC CRL 1650)
PCT-US95-02945-5

Alignment Scores:

Pred. No.: 223
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.15%
Matches: 7
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-831-458A-12 (1-325) x PCT-US95-02945-5 (1-469)

QY 47 ValProSerLeuSerGln 53

DB 483 GTCCTTATCTAACTGAA 403

RESULT 28

US-08-031-148-1
Sequence 1, Application US/08031148

Patent No. 5424398

GENERAL INFORMATION:

APPLICANT: Middelorp, Jaap Michiel.

TITLE OF INVENTION: Peptides and nucleic acid sequences

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Akzo Pharma

STREET: 1330-A Piccard Drive

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850-4377

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/031,148

FILING DATE: 19930312

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: EP 92200721.6

FILING DATE: 13-MAR-1992

ATTORNEY/AGENT INFORMATION:

NAME: Bobrowicz, Donna

REGISTRATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 538 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: double

TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:

ORGANISM: Epstein-Barr virus

US-08-031-148-1

Alignment Scores:

Pred. No.: 253
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.15%
Matches: 7
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-831-458A-12 (1-325) x US-08-041-148-1 (1-538)

QY 13 GTCCTTATCTAACTGAA 403

DB 192 GTCCTTATCTAACTGAA 403

RESULT 29

US-08-415-838-1
Sequence 1, Application US/08415838

Patent No. 6008427

GENERAL INFORMATION:

APPLICANT: Middelorp, Jaap Michiel.

TITLE OF INVENTION: Peptides and nucleic acid sequences

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Akzo No. 6008427 Patent Department

STREET: 1300 Piccard Drive, Suite 206

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/415,838

FILING DATE: 03-APR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: EP 92200721.6

FILING DATE: 13-MAR-1992

ATTORNEY/AGENT INFORMATION:

NAME: Gornley, Mary E.

REGISTRATION NUMBER: 34,409

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 538 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:

ORGANISM: Epstein-Barr virus

US-08-415-838-1

Alignment Scores:

Pred. No.: 253
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.15%
Matches: 7
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-831-458A-12 (1-325) x US-08-415-838-1 (1-538)

QY 13 GTCCTTATCTAACTGAA 403

DB 192 GTCCTTATCTAACTGAA 403

RESULT 30

US-09-205-169-1
Sequence 1, Application US/09205169

Patent No. 6365717

GENERAL INFORMATION:

APPLICANT: Middelorp, Jaap Michiel.

TITLE OF INVENTION: Peptides and nucleic acid sequences

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Akzo No. 6365717 Patent Department

STREET: 1300 Piccard Drive, Suite 206

[illegible]

```

Alignment Scores:
Pred. No.: 285 Length: 612
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.15%
Indels: 0
Gaps: 0
DB: 4

US-09-831-458A-12 (1-325) x US-09-385-982-54 (1-612)

Q7 243 LGGCHHTRCAAGSRAAH 249
DB 363 TGTGAGACTGTGAGAACAAI 373
|||||
|||||

RESULT 32
US-08-434-255-9/c
Sequence 9, Application US/08444255
Patent No. 5621089
GENERAL INFORMATION:
APPLICANT: Sloma, Alan P.
APPLICANT: Oultrup, Helge
APPLICANT: Dambmann, Claus
APPLICANT: Aaslyng, Dorrit
TITLE OF INVENTION: ALKALINE PROTEASE
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESSES:
ADDRESSER: No. 5621089/c No. 5621089/disk of No. 5621089th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-09-385-982-54
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Aulis D. ... Clark, J. H.
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 3764,400-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867-6123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 679 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-434-255-9

Alignment Scores:
Pred. No.: 315 Length: 679
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.15%
Indels: 0
Gaps: 0
DB: 1

US-09-831-458A-12 (1-325) x US-09-434-255-9 (1-679)

QY 110 GLEUHTHATQLEULYSALA 116
DB 296 GACCTGACACGATGAAAGCA 276
|||||
|||||

RESULT 33
US-08-459-967-9/c
Sequence 9, Application US/08459967

```

Patent No. 5622841
GENERAL INFORMATION:
APPLICANT: Sloma, Alan P.
APPLICANT: Outtrup, Helge
APPLICANT: Dambmann, Claus
APPLICANT: Aaslyng, Dorrit
TITLE OF INVENTION: ALKALINE PROTEASE
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESS: No. 5622841st No. 5622841st American, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.10
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,967
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/434,255
FILING DATE: 03-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Agis Dr., Cheryl H.
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 3764,400-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 679 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-459-967-9

Alignment Scores:
Pred. No.: 315
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.15%
DB: 1
Gaps: 0

US-09-831-458a-12 (1-325) x US-08-459-967-9 (1-679)

QY 110 gluleuthrArgleuLysAla 116
|||||
DB 296 GAGCTGACGATGGAAGCA 276

RESULT 34
US-08-460-327-9/c
Sequence 9, Application US/08460327
GENERAL INFORMATION:
APPLICANT: Sloma, Alan P.
APPLICANT: Outtrup, Helge
APPLICANT: Dambmann, Claus
APPLICANT: Aaslyng, Dorrit
TITLE OF INVENTION: ALKALINE PROTEASE
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESS: No. 5622841st No. 5622841st American, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA

ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.10
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,967
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/434,255
FILING DATE: 03-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Agis Dr., Cheryl H.
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 3764,400-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 679 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-460-327-9

Alignment Scores:
Pred. No.: 315
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.15%
DB: 1
Gaps: 0

US-09-831-458a-12 (1-325) x US-08-460-327-9 (1-679)

QY 110 gluleuthrArgleuLysAla 116
|||||
DB 296 GAGCTGACGATGGAAGCA 276

RESULT 35
US-08-459-871-9/c
Sequence 9, Application US/08459871
GENERAL INFORMATION:
APPLICANT: Sloma, Alan P.
APPLICANT: Outtrup, Helge
APPLICANT: Dambmann, Claus
APPLICANT: Aaslyng, Dorrit
TITLE OF INVENTION: ALKALINE PROTEASE
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESS: No. 5650326th No. 5650326th American, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.10
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,871
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/434,255
FILING DATE: 03-MAY-1995
ATTORNEY/AGENT INFORMATION:

NAME: Agria Dr., Cheryl H.
REGISTRATION NUMBER: 34,086
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 679 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-459-871-9

Alignment Scores:
Pred. No.: 315 Length: 679
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 1 Gaps: 0

US-09-831-458a-12 (1-325) x US-08-459-871-9 (1-679)

UY 110 GlutathioneS-transferase 116
DB 296 GACCTGACGATTCAGACCTT 276

RESULT 36
US-09-419-568P-24/c
Sequence 24: Application US/09419568P
Patent No. 6331613
GENERAL INFORMATION:
APPLICANT: Dumoutier, Laure
APPLICANT: Louhed, Jamila
TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible Fac
FILE REFERENCE: LUD 5543.2
CURRENT APPLICATION NUMBER: US/09/419,568P
PRIORITY FILING DATE: 1999-10-26
PRIORITY FILING DATE: 1999-07-16
PRIORITY APPLICATION NUMBER: US09/178,973
PRIORITY FILING DATE: 1998-10-26
NUMBER OF SEQ ID NOS: 29
SEQ ID NO 24
LENGTH: 690
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
US-09-419-568P-24

Alignment Scores:
Pred. No.: 319 Length: 690
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 4 Gaps: 0

US-09-831-458a-12 (1-325) x US-09-419-568P-24 (1-690)

UY 271 ScpProLeuScpProSerPhe 277
DB 550 TCTCAGCTCTCTCCACGCTT 530

RESULT 37
US-09-454-243B-24/c
Sequence 24: Application US/09454243B
Patent No. 6359117
GENERAL INFORMATION:
APPLICANT: Dumoutier, Laure

APPLICANT: Louhed, Jamila
APPLICANT: Renaud, Jean-Christophe
TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible
FILE REFERENCE: (TIFS) The Proteins Encoded, and Uses Thereof
CURRENT APPLICATION NUMBER: US/09/354,243B
PRIORITY FILING DATE: 1999-07-16
PRIORITY APPLICATION NUMBER: US09/178,973
PRIORITY FILING DATE: 1998-10-26
NUMBER OF SEQ ID NOS: 29
SEQ ID NO 24
LENGTH: 690
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
US-09-354-243B-24

Alignment Scores:
Pred. No.: 319 Length: 690
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 4 Gaps: 0

US-09-831-458a-12 (1-325) x US-09-354-243B-24 (1-690)

UY 271 ScpProLeuScpProSerPhe 277
DB 550 TCTCAGCTCTCTCCACGCTT 530

RESULT 38
US-08-458-084-4
Sequence 4: Application US/08458084
Patent No. 5624847
GENERAL INFORMATION:
APPLICANT: Fodor, William L.
APPLICANT: Hollins, Scott
APPLICANT: Squinto, Stephen P.
TITLE OF INVENTION: Chimeric Complement
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maurice M. Klee
STREET: 1951 Burr Street
CITY: Fairfield
STATE: Connecticut
COUNTRY: USA
ZIP: 06430
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 750 KB storage
COMPUTER: Dell 486/50
OPERATING SYSTEM: DOS 6.2
SOFTWARE: WordPerfect 6.0
CURRENT APPLICATION DATA:
AFFILIATION NUMBER: US/08/458,084
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Klee, Maurice M.
REGISTRATION NUMBER: 30,399
REFERENCE/DOCKET NUMBER: ALX-120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 255-1400
TELEFAX: (203) 254-1101
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 763 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Linear
MOLECULE TYPE: cDNA to mRNA

DESCRIPTION: BACIP full length cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE: Papio hamadryas
IMMEDIATE SOURCE: Baboon Spleen Lambda ZAP11 cDNA
LIBRARY: Library, Catalog # 936103,
LIBRARY: Stratagene Cloning Systems,
LIBRARY: La Jolla, California
US-08-458-084-4

Alignment Scores:
Pred. No.: 351 Length: 763
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 1 Gaps: 0

US-09-831-458a-12 (1-325) x US-08-458-084-4 (1-763)

OY 47 ValProSerLeuSerGln 53
|||||
DB 506 GTGCTTCATCCTAAGTCAA 526

RESULT 39
US-08-205-508-4
Sequence 4, Application US/08205508
Patent No. 5627264
GENERAL INFORMATION:
APPLICANT: Fodor, William L
APPLICANT: Squinto, Stephen P
TITLE OF INVENTION: Chimeric Complement
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maurice M. Klee
STREET: 1951 Burr Street
CITY: Fairfield
STATE: Connecticut
COUNTRY: USA
ZIP: 06430
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 750 Kb storage
COMPUTER: Dell 486/50
OPERATING SYSTEM: DOS 6.2
SOFTWARE: WordPerfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/205,508
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Klee, Maurice M.
REGISTRATION NUMBER: 39,399
REFERENCE/DOCKET NUMBER: ALX-120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 255-1400
TELEFAX: (203) 254-1101
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 763 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Linear
MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: BACIP full length cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE: Papio hamadryas
IMMEDIATE SOURCE: Baboon Spleen Lambda ZAP11 cDNA library, Cat
LIBRARY: # 936103, Stratagene Cloning Systems, La Jolla

LIBRARY: Baboon Spleen Lambda ZAP11 cDNA
LIBRARY: Library, Catalog # 936103,
LIBRARY: Stratagene Cloning Systems,
LIBRARY: La Jolla, California
US-08-205-508-4

Alignment Scores:
Pred. No.: 351 Length: 763
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 1 Gaps: 0

US-09-831-458a-12 (1-325) x US-08-205-508-4 (1-763)

OY 47 ValProSerLeuSerGln 53
|||||
DB 506 GTGCTTCATCCTAAGTCAA 526

RESULT 40
US-08-482-148-1
Sequence 1, Application US/08482148
Patent No. 5847082
GENERAL INFORMATION:
APPLICANT: Rother, Russell
APPLICANT: Squinto, Stephen P
TITLE OF INVENTION: Terminal Complement
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seth A. Fidel
STREET: Alexion Pharmaceuticals, 25 Science Park
CITY: New Haven
STATE: Connecticut
COUNTRY: USA
ZIP: 06511
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.4 Mb storage
COMPUTER: IBM compatible (pentium)
OPERATING SYSTEM: Windows 95 under MS DOS
SOFTWARE: Microsoft Word for Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,148
FILING DATE: 6/07/95
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Fidel, Seth A.
REGISTRATION NUMBER: 38,449
REFERENCE/DOCKET NUMBER: ALX-120,1 Div
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 776-3655
TELEFAX: (203) 776-1790
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 763 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Linear
MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: BACIP full length cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE: Papio hamadryas
IMMEDIATE SOURCE: Baboon Spleen Lambda ZAP11 cDNA library, Cat
LIBRARY: # 936103, Stratagene Cloning Systems, La Jolla
US-08-482-148-1

Alignment Scores:
Pred. No.: 351 Length: 763

Score: 7.00 Matches:
Percent Similarity: 100.00% Conservation: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 2 Gaps: 0

US-09-831-458a-12 (1-325) x US-08-482-148-1 (1-763)

OY 47 ValProSerSeuSergin 53

DB 506 GTCCTTCATCCCTAAGTCA 526

RESULT 41

PCT-US95-02944-1

Sequence 1: Application PC/TUS9502944

GENERAL INFORMATION:

APPLICANT: Kother, Russell

APPLICANT: Rollins, Scott

APPLICANT: Squinto, Stephen P

TITLE OF INVENTION: Terminal Complement

TITLE OF INVENTION: Inhibitor Fusion Genes and Proteins

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Maurice M. Klee

STREET: 1951 Burr Street

CITY: Fairfield

STATE: Connecticut

COUNTRY: USA

ZIP: 06430

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 720 Kb storage

COMPUTER: Dell 486/50

OPERATING SYSTEM: DOS 6.2

SOFTWARE: Word Perfect 6.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/02944

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/7005, 720

FILING DATE: 3-MAR-1994

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Klee, Maurice M.

REGISTRATION NUMBER: 30,399

REFERENCE/DOCKET NUMBER: ALX-129PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: (203) 255-1409

TELEFAX: (203) 254-1101

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 763 base pairs

TYPE: Nucleic Acid

STRANDEDNESS: Double

TOPOLOGY: Linear

MOLECULE TYPE: cDNA to mRNA

DESCRIPTION: BABCP full length cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Papio hamadryas

IMMEDIATE SOURCE:

LIBRARY: Baboon Spleen Lambda ZapII cDNA Library/Catalog

LIBRARY: #936103, Stratagene Cloning Systems, La Jolla, California

PCT-US95-02944-1

Alignment Scores:

Pred. No.: 351 Length: 763
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservation: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 5 Gaps: 0

US-09-831-458a-12 (1-325) x PCT US95-02944-1 (1-763)

OY 47 ValProSerSeuSergin 53

DB 506 GTCCTTCATCCCTAAGTCA 526

RESULT 42

PCT-US95-02945-4

Sequence 4: Application PC/TUS9502945

GENERAL INFORMATION:

APPLICANT: Fodor, William L

APPLICANT: Rollins, Scott

APPLICANT: Squinto, Stephen P

TITLE OF INVENTION: Chimeric Complement

TITLE OF INVENTION: Inhibitor Proteins

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: Maurice M. Klee

STREET: 1951 Burr Street

CITY: Fairfield

STATE: Connecticut

COUNTRY: USA

ZIP: 06430

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 750 Kb storage

COMPUTER: Dell 486/50

OPERATING SYSTEM: DOS 6.2

SOFTWARE: WordPerfect 6.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/02945

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/205,508

FILING DATE: 4-MAR-1994

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Klee, Maurice M.

REGISTRATION NUMBER: 30,399

REFERENCE/DOCKET NUMBER: ALX-120PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: (203) 255-1400

TELEFAX: (203) 254-1101

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 763 base pairs

TYPE: Nucleic Acid

STRANDEDNESS: Double

TOPOLOGY: Linear

MOLECULE TYPE: cDNA to mRNA

DESCRIPTION: BABCP full length cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Papio hamadryas

IMMEDIATE SOURCE:

LIBRARY: Baboon Spleen Lambda ZapII cDNA Library/Catalog

LIBRARY: #936103, Stratagene Cloning Systems, La Jolla, California

PCT-US95-02945-4

Alignment Scores:

Pred. No.: 351 Length: 763
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservation: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 5 Gaps: 0

US-09-831-458a-12 (1-325) x PCT US95-02945-4 (1-763)

OY 47 ValProSerSeuSergin 53

DB 506 GTCCTTCATCCCTAAGTCA 526

Db 506 GTCGATGAGCTTCCTTCC 525

RESULT 43

US-09-051-696-5/c

Sequence 5, Application US/09051696

Patent No. 6319495

GENERAL INFORMATION:

APPLICANT: Pollock, Allan S.
Christakos, Sylvia
Keddy, Daphne

TITLE OF INVENTION: Method for Restoring Glucose Responsiveness to Insulin Secretion

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Crow LLP
STREET: Two Embarcadero Center, Fifth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/051,696

FILING DATE: 17-Apr-1998

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/005,386

FILING DATE: 10-OCT-1995

APPLICATION NUMBER: WO PCT/US96/16736

FILING DATE: 17-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677

REFERENCE/DOCKET NUMBER: 02307E-062710US

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 791 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-09-051-696-5

Alignment Scores:

Pred. No.: 363

Score: 7.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 2.15%

Matches: 791

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

US-09-831-458a-12 (1-325) x US-09-051-696-5 (1-791)

Db 119 GTCGACGAGCTTCCTTCC 99

RESULT 44

US-09-416-509C-2/c

Sequence 2, Application US/09416509C

Patent No. 6469141

GENERAL INFORMATION:

APPLICANT: Bergson, Claire

TITLE OF INVENTION: D1-like Dopamine Receptor Activity Modifying Protein

FILE REFERENCE: M0351-205012

US-09-831-458a-12 (1-325) x US-09-416-509C-2 (1-936)

Db 268 ValAspClySerProLeuSer 274

Db 115 GTCGATGAGCTTCCTTCC 99

RESULT 45

US-08-973-068-28

Sequence 28, Application US/08973068

Patent No. 6127604

GENERAL INFORMATION:

APPLICANT: Dale, James Langham

APPLICANT: Harding, Robert Maxwell

APPLICANT: Dugdale, Benjamin

APPLICANT: Beetham, Peter Ronald

APPLICANT: Haefel, Gregory John

TITLE OF INVENTION: INTERGENIC REGIONS OF BANANA BUNNY

FILE REFERENCE: 09657/002001

CURRENT APPLICATION NUMBER: US/08/973,068

CURRENT FILING DATE: 1998-03-12

EARLIER APPLICATION NUMBER: PCT/AU96/00445

EARLIER FILING DATE: 1996-05-11

NUMBER OF SEQ ID NOS: 61

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 28

LENGTH: 982

TYPE: DNA

ORGANISM: Banana Bunchy Top Virus (BBTV)

US-08-973-068-28

Alignment Scores:

Pred. No.: 444

Score: 7.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 2.15%

Matches: 982

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

US-09-831-458a-12 (1-325) x US-08-973-068-28 (1-982)

Db 941 TCGAAGTCCTTCGAGTTT 961

Db 45 SerIysValProSerSerIen 51

Search completed: December 7, 2002, 14:54:28

Job time : 81 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Computer 1.13

OM protein - nucleic search, using frame_plus_f2n model

Run on: December 7, 2002, 12:27:12, Search time: 61 seconds

(without alignments)
2077.510 Million cell updates/sec

Title: US-09-831-458A-12

Perfect score: 325
Sequence: 1 MSOSKEPRVQGLIIPGLGH PCTVNWYVCKKPAKFPNP 325

Scoring table:

OLIGO
Xgapop 60.0, Ygapext 60.0
Ygapop 60.0, Ygapext 60.0
Delop 6.0, Delext 7.0

Searched: 350425 seqs, 19466365 residues

Word size: 5

Total number of hits satisfying chosen parameters: 43445

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_f2n_model -DEV=xib
-O=/cgn2_1/USPTO/US09831458/rmat.05122002.1041416803/ffp_juncy.fasta_1_113
-DB=Published.Applications_USA_QFMT=fastap -SUFFIX=ol1.rmpb -MINMATCH=0 1
-LOOKUP=0 -LOOKEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=ol1go
-TRANS-human40.cdi -LIST=45 -LOCALIGN=200 -THR_SCORE=quality -THR_MIN=5
-ALIGN=45 -MODE=LOCAL -OUTEXT=ptc -NORM=ext -HEAPS17E=500 -MINFN=0
-MAXLEN=200000000 -USER=US09831458 -ACGN.1.1.21.tunat.05122002.1041416803
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MAP -LARGOUDRY -NCG_SCORES=0 -WAIT -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database: Published.Applications_USA:

1: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubppa/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.seq:
5: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubppa/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.seq:
9: /cgn2_6/ptodata/2/pubppa/US09_NEW_PUB.seq:
10: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.seq:
12: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.seq:
13: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.seq:
14: /cgn2_6/ptodata/2/pubppa/US06_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	15.7	592	US-09-864-761-7350	Sequence 7350, Ap
2	50	15.4	152	US-09-864-761-24072	Sequence 24072, A
3	9	2.8	1183	US-09-983-531A-15	Sequence 15, Appl
4	9	2.8	2404	US-09-983-531A-25	Sequence 25, Appl

5	9	2.8	2669	US-09-983-531A-19	Appl
6	9	2.8	2757	US-10-020-743-2	Appl
7	9	2.8	2757	US-10-125-470-2	Appl
8	9	2.8	2757	US-10-125-452-2	Appl
9	9	2.8	2781	US-10-020-743-1	Appl
10	9	2.8	2868	US-10-020-743-7	Appl
11	9	2.8	2892	US-10-020-743-5	Appl
12	9	2.8	3512	US-10-020-743-9	Appl
13	8	2.5	623	US-09-894-638A-110	Appl
14	8	2.5	884	US-09-770-445-515	Appl
15	8	2.5	893	US-09-894-638A-95	Appl
16	8	2.5	952	US-09-770-445-328	Appl
17	8	2.5	1149	US-09-887-576-673	Appl
18	8	2.5	2006	US-09-887-576-826	Appl
19	8	2.5	3150	US-09-962-436-331	Appl
20	8	2.5	3150	US-09-969-708-453	Appl
21	8	2.5	3150	US-09-954-456-1224	Appl
22	8	2.5	18871	US-09-764-847-1317	Appl
23	7	2.2	124	US-09-294-0938-2487	Appl
24	7	2.2	128	US-10-046-935-258	Appl
25	7	2.2	128	US-09-879-178-258	Appl
26	7	2.2	238	US-09-864-761-23899	Appl
27	7	2.2	258	US-09-923-876-704	Appl
28	7	2.2	258	US-09-923-876-4835	Appl
29	7	2.2	258	US-09-923-876-4835	Appl
30	7	2.2	275	US-09-923-876-4409	Appl
31	7	2.2	275	US-09-878-574-13627	Appl
32	7	2.2	314	US-09-764-865-2051	Appl
33	7	2.2	314	US-09-764-865-2052	Appl
34	7	2.2	326	US-09-983-965-1896	Appl
35	7	2.2	336	US-09-924-300-2012	Appl
36	7	2.2	342	US-09-810-997-8	Appl
37	7	2.2	342	US-09-880-371-14	Appl
38	7	2.2	342	US-09-880-371-15	Appl
39	7	2.2	342	US-09-880-371-15	Appl
40	7	2.2	349	US-09-983-965-2185	Appl
41	7	2.2	354	US-09-867-701-3187	Appl
42	7	2.2	363	US-09-880-107-182	Appl
43	7	2.2	364	US-09-879-574-182	Appl
44	7	2.2	384	US-09-960-352-2949	Appl
45	7	2.2	408	US-09-829-124-6	Appl

ALIGNMENTS

RESULT 1
US-09-864-761-7350
Sequence 7350, Application US/09864761
Patent No US/20048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Hanzel, David R.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLE
FILE REFERENCE: Accoml-X-1
CURRENT APPLICATION NUMBER: US-09-974-761
CURRENT FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: US 60/180,412
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/227,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/642,466
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 2426436
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/246,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30

[illegible]

```

Db      63 TCTACTCCGAGCGGAATGCTACACTGCTCTACCTGAGTGAAGTGAAGCCGAC 122
      63 TCTACTCCGAGCGGAATGCTACACTGCTCTACCTGAGTGAAGTGAAGCCGAC 122
OY      229 LeuValValIleLysThrAlaGluGlu 238
      123 GTCGTGTAATCAAAACTGCTGAGAGAGAG 152

RESULT 3
US-09-983-531A-15
: Sequence 15, Application US/09983531A
: Patent No. US20020147132A1
: GENERAL INFORMATION:
: APPLICANT: Fujisawa, Atsuko
: APPLICANT: Yamakawa, Toru
: APPLICANT: Shirakawa, Kamon
: APPLICANT: Chitose, Oriti
: APPLICANT: Ogawa, Naoki
: TITLE OF INVENTION: Meltrins
: FILE REFERENCE: 11-22-99 sequence submission
: CURRENT APPLICATION NUMBER: US/09/983,531A
: PRIORITY FILING DATE: 2001-10-24
: PRIOR APPLICATION NUMBER: JP 8-61756
: PRIOR FILING DATE: 1996-02-23
: PRIOR APPLICATION NUMBER: PCT/JP96/03017
: PRIOR FILING DATE: 1996-10-17
: NUMBER OF SEQ ID NOS: 44
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 15
: LENGTH: 1183
: TYPE: DNA
: ORGANISM: Unknown organism
: FEATURE:
: OTHER INFORMATION: Description of Unknown Organism: Human meltrin
: NAME/KEY: CDS
: LOCATION: (2)..(1183)
: US-09-983-531A-15

Alignment Scores:
Pred. No.:      3.22      Length:      1183
Score:          9.00      Matches:      9
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      2.77%      Indels:      0
DB:              10      Gaps:      0

US-09-831-458A-12 (1-325) x US-09-983-531A-15 (1-1183)
OY      34 AlaGlyValLeuValAlaIleLeuVal 42
      998 GGTGAGTGTGTGTCGTCATCTTGTG 1024

RESULT 4
US-09-983-531A-25
: Sequence 25, Application US/09983531A
: Patent No. US20020147132A1
: GENERAL INFORMATION:
: APPLICANT: Fujisawa, Atsuko
: APPLICANT: Yamakawa, Toru
: APPLICANT: Shirakawa, Kamon
: APPLICANT: Chitose, Oriti
: APPLICANT: Ogawa, Naoki
: TITLE OF INVENTION: Meltrins
: FILE REFERENCE: 11-22-99 sequence submission
: CURRENT APPLICATION NUMBER: US/09/983,531A
: PRIORITY FILING DATE: 2001-10-24
: PRIOR APPLICATION NUMBER: JP 8-61756
: PRIOR FILING DATE: 1996-02-23
: PRIOR APPLICATION NUMBER: PCT/JP96/03017
: PRIOR FILING DATE: 1996-10-17
: NUMBER OF SEQ ID NOS: 44
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 25

```

```

: LENGTH: 2404
: TYPE: DNA
: ORGANISM: Unknown organism
: FEATURE:
: OTHER INFORMATION: Description of Unknown Organism: (1)
: NAME/KEY: CDS
: LOCATION: (1)..(1017)
: US-09-983-531A-25

Alignment Scores:
Pred. No.:      6.38      Length:      2404
Score:          9.00      Matches:      9
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      2.77%      Indels:      0
DB:              10      Gaps:      0

US-09-831-458A-12 (1-325) x US-09-983-531A-25 (1-2404)
OY      34 AlaGlyValLeuValAlaIleLeuVal 42
      733 GCTGAGTGTGTGTCGTCATCTTGTG 759

RESULT 5
US-09-983-531A-19
: Sequence 19, Application US/09983531A
: Patent No. US20020147132A1
: GENERAL INFORMATION:
: APPLICANT: Fujisawa, Atsuko
: APPLICANT: Yamakawa, Toru
: APPLICANT: Shirakawa, Kamon
: APPLICANT: Chitose, Oriti
: APPLICANT: Ogawa, Naoki
: TITLE OF INVENTION: Meltrins
: FILE REFERENCE: 11-22-99 sequence submission
: CURRENT APPLICATION NUMBER: US/09/983,531A
: PRIORITY FILING DATE: 2001-10-24
: PRIOR APPLICATION NUMBER: JP 8-61756
: PRIOR FILING DATE: 1996-02-23
: PRIOR APPLICATION NUMBER: PCT/JP96/03017
: PRIOR FILING DATE: 1996-10-17
: NUMBER OF SEQ ID NOS: 44
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 19
: LENGTH: 2669
: TYPE: DNA
: ORGANISM: Unknown organism
: FEATURE:
: OTHER INFORMATION: Description of Unknown Organism: (1)
: OTHER INFORMATION: JMI09(pmel beta 240) and JMI09(pmel beta
: NAME/KEY: CDS
: LOCATION: (2)..(1282)
: US-09-983-531A-19

Alignment Scores:
Pred. No.:      7.06      Length:      2669
Score:          9.00      Matches:      9
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      2.77%      Indels:      0
DB:              10      Gaps:      0

US-09-831-458A-12 (1-325) x US-09-983-531A-19 (1-2669)
OY      34 AlaGlyValLeuValAlaIleLeuVal 42
      998 GGTGAGTGTGTGTCGTCATCTTGTG 1024

RESULT 6
US-10-020-733-3
: Sequence 3, Application US/10020733

```

```

Patent No. US20020161214A1
GENERAL INFORMATION:
APPLICANT: Waikie, D. Wade
APPLICANT: Wilgowski, Nathaniel L.
APPLICANT: Friddele, Carl Johan
TITLE OF INVENTION: No. US20020161214A1-1 Human Proteases and Polypeptides
FILE REFERENCE: LEX-0263-USA
CURRENT APPLICATION NUMBER: US/10/020,733
CURRENT FILING DATE: 2001-10-30
PRIOR APPLICATION NUMBER: US 60/244,939
PRIOR FILING DATE: 2000-11-01
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2757
TYPE: DNA
ORGANISM: homo sapiens
US-10-020-733-3

```

```

Alignment Scores:
Pred. No.: 7 29 Length: 2757
Score: 9 00 Matches: 0
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2 77% Indels: 0
DB: 9 Gaps: 0

```

```

US-09-831-458A-12 (1-325) x US-10-020-733-3 (1-2757)

```

```

Q7 34 AlaclyValLeuValAlaIleLeuVal 42
|||||
DB 2113 GCTGGAGTCTCTCTGCGCATCTTCGCTG 2139

```

```

RESULT 7
US-10-125-470-2
Sequence 2, Application US/10125470
Patent No. US20020165377A1
GENERAL INFORMATION:
APPLICANT: Popen et al.
TITLE OF INVENTION: ADAM Polynucleotides, polypeptides, and Antibodies
FILE REFERENCE: PRO06P1
CURRENT APPLICATION NUMBER: US/10/125,470
CURRENT FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: US/09/712,907A
PRIOR FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: 60/136,488
PRIOR FILING DATE: 1999-05-27
PRIOR APPLICATION NUMBER: 60/142,930
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 60/178,717
PRIOR FILING DATE: 2000-01-28
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 2757
TYPE: DNA
ORGANISM: Homo sapiens
US-10-125-470-2

```

```

Alignment Scores:
Pred. No.: 7 29 Length: 2757
Score: 9 00 Matches: 0
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2 77% Indels: 0
DB: 9 Gaps: 0

```

```

US-09-831-458A-12 (1-325) x US-10-125-470-2 (1-2757)

```

```

Q7 34 AlaclyValLeuValAlaIleLeuVal 42
|||||
DB 2113 GCTGGAGTCTCTCTGCGCATCTTCGCTG 2139

```

```

RESULT 8
US-10-125-452-2
Sequence 2, Application US/10125452
Patent No. US20020173640A1
GENERAL INFORMATION:
APPLICANT: Edeh et al.
TITLE OF INVENTION: ADAM Polynucleotides, polypeptides, and Antibodies
FILE REFERENCE: PRO06P2
CURRENT APPLICATION NUMBER: US/10/125,452
CURRENT FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/155,504
PRIOR FILING DATE: 2001-09-19
PRIOR APPLICATION NUMBER: 60/172,907
PRIOR FILING DATE: 2000-11-16
PRIOR APPLICATION NUMBER: 60/180,14308
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: 60/178,717
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: 60/142,930
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 60/136,488
PRIOR FILING DATE: 1999-05-27
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 2757
TYPE: DNA
ORGANISM: Homo sapiens
US-10-125-452-2

```

```

Alignment Scores:
Pred. No.: 7 29 Length: 2757
Score: 9 00 Matches: 0
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2 77% Indels: 0
DB: 9 Gaps: 0

```

```

US-09-831-458A-12 (1-325) x US-10-125-452-2 (1-2757)

```

```

Q7 34 AlaclyValLeuValAlaIleLeuVal 42
|||||
DB 2113 GCTGGAGTCTCTCTGCGCATCTTCGCTG 2139

```

```

RESULT 9
US-10-020-733-1
Sequence 1, Application US/10020733
Patent No. US20020161214A1
GENERAL INFORMATION:
APPLICANT: Waikie, D. Wade
APPLICANT: Wilgowski, Nathaniel L.
APPLICANT: Friddele, Carl Johan
TITLE OF INVENTION: No. US20020161214A1-1 Human Proteases and Polypeptides
FILE REFERENCE: LEX-0263-USA
CURRENT APPLICATION NUMBER: US/10/020,733
CURRENT FILING DATE: 2001-10-30
PRIOR APPLICATION NUMBER: US 60/244,939
PRIOR FILING DATE: 2000-11-01
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2781
TYPE: DNA
ORGANISM: Homo sapiens
US-10-020-733-1

```

```

Alignment Scores:
Pred. No.: 7 35 Length: 2781
Score: 9 00 Matches: 0
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2 77% Indels: 0
DB: 9 Gaps: 0

```

US-09-831-458A-12 (1-325) x US-10-020-733-1 (1-2781)

QY 34 AlaglyValleuValAla1leuVal 42

DB 2113 GCTGAGCTGTGTCGCATCTTCATG 2139

RESULT 10

US-10-020-733-7

Sequence 7, Application US/10020733

Patent No. US20020161214A1

GENERAL INFORMATION:

APPLICANT: Walke, D. Wade

APPLICANT: Wilganowski, Nathaniel L.

TITLE OF INVENTION: No. US20020161214A1 Human Proteases and Polypeptides Encodin

FILE REFERENCE: LEX-0263-USA

CURRENT APPLICATION NUMBER: US/10/020,733

PRIOR FILING DATE: 2001-10-30

PRIOR APPLICATION NUMBER: US 60/244,939

PRIOR FILING DATE: 2000-11-01

NUMBER OF SEQ ID NOS: 9

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7

LENGTH: 2868

TYPE: DNA

ORGANISM: homo sapiens

US-10-020-733-7

Alignment Scores:

Pred. No.: 7.57

Score: 9.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 2.77%

DB: 9

Matches: 2868

Conservative: 9

Mismatches: 0

Indels: 0

Gaps: 0

US-09-831-458A-12 (1-325) x US-10-020-733-7 (1-2868)

QY 34 AlaglyValleuValAla1leuVal 42

DB 2113 GCTGAGCTGTGTCGCATCTTCATG 2139

RESULT 11

US-10-020-733-5

Sequence 5, Application US/10020733

Patent No. US20020161214A1

GENERAL INFORMATION:

APPLICANT: Walke, D. Wade

APPLICANT: Wilganowski, Nathaniel L.

TITLE OF INVENTION: No. US20020161214A1 Human Proteases and Polypeptides Encodin

FILE REFERENCE: LEX-0263-USA

CURRENT APPLICATION NUMBER: US/10/020,733

PRIOR FILING DATE: 2001-10-30

PRIOR APPLICATION NUMBER: US 60/244,939

PRIOR FILING DATE: 2000-11-01

NUMBER OF SEQ ID NOS: 9

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 5

LENGTH: 2892

TYPE: DNA

ORGANISM: homo sapiens

US-10-020-733-5

Alignment Scores:

Pred. No.: 7.64

Score: 9.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 2.77%

DB: 9

Matches: 2892

Conservative: 9

Mismatches: 0

Indels: 0

Gaps: 0

US-09-831-458A-12 (1-325) x US-10-020-733-9 (1-2892)

QY 34 AlaglyValleuValAla1leuVal 42

DB 2113 GCTGAGCTGTGTCGCATCTTCATG 2139

RESULT 12

US-10-020-733-9

Sequence 9, Application US/10020733

Patent No. US20020161214A1

GENERAL INFORMATION:

APPLICANT: Walke, D. Wade

APPLICANT: Wilganowski, Nathaniel L.

TITLE OF INVENTION: No. US20020161214A1 Human Proteases

FILE REFERENCE: LEX-0263-USA

CURRENT APPLICATION NUMBER: US/10/020,733

PRIOR FILING DATE: 2001-10-30

PRIOR APPLICATION NUMBER: US 60/244,939

PRIOR FILING DATE: 2000-11-01

NUMBER OF SEQ ID NOS: 9

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 9

LENGTH: 3512

TYPE: DNA

ORGANISM: homo sapiens

US-10-020-733-9

Alignment Scores:

Pred. No.: 5.2

Score: 9.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 2.77%

DB: 9

Matches: 3512

Conservative: 9

Mismatches: 0

Indels: 0

Gaps: 0

US-09-831-458A-12 (1-325) x US-10-020-733-9 (1-3512)

QY 34 AlaglyValleuValAla1leuVal 42

DB 2335 GCTGAGCTGTGTCGCATCTTCATG 2461

RESULT 13

US-09-894-633A-110/C

Sequence 110, Application US/09894633A

Patent No. US20020124285A1

GENERAL INFORMATION:

APPLICANT: Comer, Timothy

APPLICANT: Dubois, Patricia

APPLICANT: Maiver, Marianne

APPLICANT: Masucci, James

TITLE OF INVENTION: PLANT BIOLOGICAL AGENTS FOR SELECT

FILE REFERENCE: 38-21(15856)B

CURRENT APPLICATION NUMBER: US/09/894,633A

PRIOR FILING DATE: 2001-06-28

PRIOR APPLICATION NUMBER: 60/214,457

PRIOR FILING DATE: 2000-06-28

PRIOR APPLICATION NUMBER: 09/894,633

PRIOR FILING DATE: 2000-06-28

NUMBER OF SEQ ID NOS: 111

SOFTWARE: PatentIn version 4.0

SEQ ID NO 110

LENGTH: 623

TYPE: DNA

ORGANISM: Zea mays

US-09-894-633A-110

Alignment Scores:

Pred. No.: 18.7

Score: 8.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 2.46%

DB: 623

Matches: 623

Conservative: 8

Mismatches: 0

Indels: 0

Gaps: 0

Percent Similarity:	100.00%	Conservative:	0
Best local Similarity:	100.00%	Mismatches:	0
Query Match:	2.46%	Indels:	0
IN:	10	Gaps:	0

```

CURRENT FILING DATE: 2001-10-03
PRIOR APPLICATION NUMBER: US/60/237,606
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: US/60/237,606
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: US/60/237,625
PRIOR FILING DATE: 2000-10-03
NUMBER OF SEQ ID NOS: 658
SOFTWARE: patentin version 3.0
SEQ ID NO 453
LENGTH: 4150
TYPE: DNA
ORGANISM: Homo sapiens
US-09-969-708-453

Alignment Scores:
Pred. No.: 89
Score: 8.00
Length: 3150
Percent similarity: 100.00%
Matches: 8
Best Local Similarity: 100.00%
Conservative: 0
MisMatches: 0
Query Match: 2.46%
Indels: 0
Gaps: 0
DB: 10

US-09-831-458a-12 (1-325) x US-09-969-708-453 (1-3150)
CY 38 ValAla111eUuValGInuAIser 45
|||||
DB 1089 GTACCTCATCTTCGTCACAGCTTTC 1086

RESULT 21
US-09-954-456-1224/c
Sequence 1224, Application US/99/44456
Patent No. US2001015057A1
GENERAL INFORMATION:
APPLICANT: Young, Paul
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
FILE REFERENCE: 689290-76
CURRENT APPLICATION NUMBER: US/60/234,456
CURRENT FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/60/233,617
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,052
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,123
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,134
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,637
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,638
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,711
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,720
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,840
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,864
PRIOR FILING DATE: 2000-09-27
NUMBER OF SEQ ID NOS: 2276
SOFTWARE: patentin version 3.0
SEQ ID NO 1224
LENGTH: 4150
TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-456-1224

Alignment Scores:
Pred. No.: 89
Score: 8.00
Length: 3150
Percent similarity: 100.00%
Matches: 8
Best Local Similarity: 100.00%
Conservative: 0
MisMatches: 0
Query Match: 2.46%
Indels: 0
Gaps: 0
DB: 10

```

```

Query Match: 2.46%
Indels: 0
Gaps: 0
DB: 10

US-09-831-458a-12 (1-325) x US-09-954-456-1224 (1-3150)
CY 39 ValAla111eUuValGInuAIser 45
|||||
DB 1089 GTACCTCATCTTCGTCACAGCTTTC 1086

RESULT 22
US-09-764-847-1317/c
Sequence 1317, Application US/97/64847
Patent No. US20020132767A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: 687004
CURRENT APPLICATION NUMBER: US/99/764,847
CURRENT FILING DATE: 2001-01-17
PRIOR APPLICATION data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 2004
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 1317
LENGTH: 18871
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-847-1317

Alignment Scores:
Pred. No.: 499
Score: 8.00
Length: 18871
Percent similarity: 100.00%
Matches: 8
Best Local Similarity: 100.00%
Conservative: 0
MisMatches: 0
Query Match: 2.46%
Indels: 0
Gaps: 0
DB: 10

US-09-831-458a-12 (1-325) x US-09-764-847-1317 (1-18871)
CY 44 ValSec15ValProGserGser 51
|||||
DB 1984 GTTCGACGCTTCATCCAGCTTTC 1961

RESULT 23
US-09-294-093b-2487/c
Sequence 2487, Application US/97/4093b
Patent No. US20010051335A1
GENERAL INFORMATION:
APPLICANT: Ito, Taisandi, Kashinath, V.
APPLICANT: Ito, Laura, Y.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
FILE REFERENCE: PL-0003 US
CURRENT APPLICATION NUMBER: US/99/294,093b
CURRENT FILING DATE: 1999-04-16
PRIOR APPLICATION NUMBER: US/98/242,527
PRIOR FILING DATE: April 21, 1998
NUMBER OF SEQ ID NOS: 6207
SOFTWARE: PERL Program
SEQ ID NO 2487
LENGTH: 124
TYPE: DNA
ORGANISM: Zoa mays
FEATURE:
NAME/KEY: misc. feature
OTHER INFORMATION: Inocyte ID No. US20010051335A1 700346184H1
NAME/KEY: unsure
LOCATION: 5', 40, 59, 73
OTHER INFORMATION: a, c, g, or other
US-09-294-093b-2487

Alignment Scores:
Pred. No.: 42.4
Score: 7.00
Length: 124
Matches: 7

```


Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.15%
DB: 10
Conservative: 0
Matches: 0
Indels: 0
Gaps: 0

US-09-831-458a-12 (1-325) x US-09-204-0938 2487 (1-124)

OY 35 GlyValLeuValAlaIleLeu 41
DB 21 GCGCTGCTGTGCAATAC 1

RESULT 24

US-10-046-935-1658
Sequence 1658, Application US/10046935
Patent No. US2002015601A1
GENERAL INFORMATION:
APPLICANT: Jiang, Yujin
APPLICANT: Harlocker, Susan L.
APPLICANT: Secrist, Heather
APPLICANT: Wang, Aljun
APPLICANT: Stolk, John A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.52721
CURRENT APPLICATION NUMBER: US/10/046,935
CURRENT FILING DATE: 2002-01-15
NUMBER OF SEQ. ID NOS: 2239
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1658
LENGTH: 128
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9, 10, 11, 53, 68
OTHER INFORMATION: n = A,T,C or G
US-10-046-935-1658

Alignment Scores:

Prod. No.: 43.7
Score: 7.00
Length: 128
Matches: 7
Conservative: 0
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.15%
Indels: 0
Gaps: 0

US-09-831-458a-12 (1-325) x US-10-046-935-1658 (1-128)

OY 10 GlnGlnLeuGlyLeuLeuGly 16
DB 104 CAGCAACTGGCTTACTAGG 124

RESULT 25

US-09-878-178-1658
Sequence 1658, Application US/09878178
Patent No. US20020177552A1
GENERAL INFORMATION:
APPLICANT: Jiang, Yujin
APPLICANT: Harlocker, Susan L.
APPLICANT: Secrist, Heather
APPLICANT: Wang, Aljun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.527
CURRENT APPLICATION NUMBER: US/09/878,178
CURRENT FILING DATE: 2001-06-08
NUMBER OF SEQ. ID NOS: 2237
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1658
LENGTH: 128
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature

LOCATION: (1)-(128)
OTHER INFORMATION: n = A,T,C or G
US-09-878-178-1658

Alignment Scores:

Prod. No.: 43.7
Score: 7.00
Length: 128
Matches: 7
Conservative: 0
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.15%
Indels: 0
Gaps: 0

US-09-831-458a-12 (1-325) x US-09-878-178-1658 (1-128)

OY 10 GlnGlnLeuGlyLeuLeuGly 16
DB 104 CAGCAACTGGCTTACTAGG 124

RESULT 26

US-09-864-761-23399/c
Sequence 23399, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David R.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME DERIVED SINGLE EXON NUCLE
FILE REFERENCE: Acomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,412
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/654,466
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: GR 242644
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/226,459
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US00/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/244,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ. ID NOS: 49117
SOFTWARE: Anomax Sequence List and Editing vers. 1.1
SEQ ID NO 23399
LENGTH: 238
TYPE: DNA

```

US 09-8341458A 12 (1-325), X US-09-923-876-704 (1-258)
QY 285 GluProAsnAsnSerGlyAsn 291
      |||||||
DD 14 GAGCTTAACACACAGTGGAAAC 34

```

```

? ORGANISM: Zoa myx
?
? FEATURE:
? NAME/KEY: misc10feature_19 NC_018200.20013058A1.79045585A
? OTHER INFORMATION: unsure
? NAME/KEY: unsure
? LOCATION: 40, 46, 186, 222
? OTHER INFORMATION: a, c, g, t, or other
? OS-09-923-876 4826

```

Alignment Scores:	89.1	Length:	268
Pred. No.:	7.00	Matches:	7
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:			

```

Query Match: 2.15% Indels: 0
DB: 10 Gaps: 0
US-09-831-458a-12 (1-325) x US-09-923-876-4825 (1-268)

QY 35 G1yValLeuValAlaIleLeu 41
DB 123 GCGGTGCTTTCATGATCTG 103

RESULT 30
US-09-923-876-4499/c
: Sequence 4499, Application US/09924876
: Patent No. US20020013958A1
: GENERAL INFORMATION:
: APPLICANT: Lalupdi, Raghunath V.
: APPLICANT: Kamigaki, Laura Y. (ILO)
: APPLICANT: Sherman, Bradley K.
: TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
: FILE REFERENCE: PI-0012-1-COR
: CURRENT APPLICATION NUMBER: US/09/924,876
: PRIOR FILING DATE: 2001-08-06
: PRIOR APPLICATION NUMBER: 09/298,329
: PRIOR FILING DATE: 1999-04-21
: PRIOR APPLICATION NUMBER: 60/085,331
: PRIOR FILING DATE: 1998-05-05
: NUMBER OF SEQ ID NOS: 4332
: SOFTWARE: PERL Program
: SEQ ID NO 4499
: LENGTH: 275
: TYPE: DNA
: ORGANISM: Zea mays
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Intron ID No. US20020013958A1 700455143H1
: LOCATION: 257
: OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-4499

Alignment Scores:
Pred. No.: 91.3 Length: 275
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 10 Gaps: 0
US-09-831-458a-12 (1-325) x US-09-923-876-4499 (1-275)

QY 35 G1yValLeuValAlaIleLeu 41
DB 127 GCGGTGCTTTCATGATCTG 107

RESULT 31
US-09-878-574-13627/c
: Sequence 13627, Application US/09878574
: Patent No. US20020110548A1
: GENERAL INFORMATION:
: APPLICANT: Byrum, Joseph R.
: APPLICANT: La Kosa, Thomas J.
: APPLICANT: Thompson, Michael D.
: TITLE OF INVENTION: Nucleic Acid Molecules and other Molecules Associated with
: FILE REFERENCE: 38-21(15401)B
: CURRENT APPLICATION NUMBER: US/09/878,574
: PRIOR FILING DATE: 2001-12-21
: PRIOR APPLICATION NUMBER: 09/333,535
: PRIOR FILING DATE: 1999-06-14
: NUMBER OF SEQ ID NOS: 15775
: SEQ ID NO 13627
: LENGTH: 275
: TYPE: DNA
: ORGANISM: Glycine max

```

```

: OTHER INFORMATION: Clone ID: 701067461H1
US-09-878-574-13627

Alignment Scores:
Pred. No.: 91.3 Length: 275
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 10 Gaps: 0
US-09-831-458a-12 (1-325) x US-09-878-574-13627 (1-275)

QY 271 SerProLeuSerProSerPhe 277
DB 21 TCGGCACCTTCACATCTTTC 1

RESULT 32
US-09-764-869-2051
: Sequence 2051, Application US/09764869
: Patent No. US20020061521A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PC007
: CURRENT APPLICATION NUMBER: US/09/764,869
: PRIOR FILING DATE: 2001-01-17
: PRIOR APPLICATION data removed - refer to PALM or file who
: NUMBER OF SEQ ID NOS: 2442
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2051
: LENGTH: 314
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-764-869-2051

Alignment Scores:
Pred. No.: 104 Length: 314
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 10 Gaps: 0
US-09-831-458a-12 (1-325) x US-09-764-869-2051 (1-314)

QY 27 LouLeuSerPheMetLeuLeu 33
DB 126 CTCCTTCTTCATGATCTGCTG 116

RESULT 33
US-09-764-869-2052
: Sequence 2052, Application US/09764869
: Patent No. US20020061521A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PC007
: CURRENT APPLICATION NUMBER: US/09/764,869
: PRIOR FILING DATE: 2001-01-17
: PRIOR APPLICATION data removed - refer to PALM or file who
: NUMBER OF SEQ ID NOS: 2442
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2052
: LENGTH: 314
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-764-869-2052

Alignment Scores:
Pred. No.: 104 Length: 314
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0

```

Best Local Similarity: 100.00%
 Query Match: 2.15%
 DB: 10
 Matches: 0
 Indels: 0
 Gaps: 0

US-09-831-458a-12 (1-325) x US-09-764-869-2052 (1-314)

QY 27 LeuleusPhemeLeuLeu 43
 DB 126 CTCTTCTTCATGTTGCTG 146

RESULT 34

US-09-983-965-1896

Sequence 1896, Application US/0994365

Patent No. US20020137160A1

GENERAL INFORMATION:

APPLICANT: Warren, Wesley C.

APPLICANT: Tao, Ningbing

APPLICANT: Hyatt, John C.

APPLICANT: Mathalagan, Nagappan

TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LATENTATION AND

FILE REFERENCE: 37-21(10297)C

CURRENT APPLICATION NUMBER: US/09/983,965

CURRENT FILING DATE: 2001-10-26

PRIOR APPLICATION NUMBER: US 09/465,231

PRIOR FILING DATE: 1999-12-15

PRIOR APPLICATION NUMBER: US 60/113,678

PRIOR FILING DATE: 1998-12-17

NUMBER OF SEQ ID NOS: 5912

SEQ ID NO 1896

LENGTH: 326

TYPE: DNA

ORGANISM: Bos taurus

FEATURE:

OTHER INFORMATION: Clone ID: 11-L1H3057-007-Q1-K1-C3

US-09-983-965-1896

Alignment Scores:

Pred. No.: 108 Length: 326

Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.15% Indels: 0

DB: 10 Gaps: 0

US-09-831-458a-12 (1-325) x US-09-983-965-1896 (1-326)

QY 24 ValleusLeuLeuSerPhe 30

DB 106 GTTTTAAAGCTTCTGAGTTT 126

RESULT 35

US-09-974-300-2011/C

Sequence 2011, Application US/09974300

Patent No. US20020146721A1

GENERAL INFORMATION:

APPLICANT: Berka, Randy M.

APPLICANT: Clausen, Ib Groth

TITLE OF INVENTION: Methods For Monitoring Multiple Gene

FILE REFERENCE: 10085,500-US

CURRENT APPLICATION NUMBER: US/09/974,300

CURRENT FILING DATE: 2001-10-05

PRIOR APPLICATION NUMBER: 09/640,548

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: 60/279,526

PRIOR FILING DATE: 2001-03-27

NUMBER OF SEQ ID NOS: 8481

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 2011

LENGTH: 346

TYPE: DNA

ORGANISM: Bacillus licheniformis

FEATURE:

OTHER INFORMATION:

US-09-974-300-2011

Alignment Scores:

Pred. No.: 111 Length: 336

Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.15% Indels: 0

DB: 10 Gaps: 0

US-09-831-458a-12 (1-325) x US-09-974-300-2011 (1-336)

QY 23 LeCValleusLeuLeuSer 29

DB 22 TTAACTTTCAGCTCTTCA 2

RESULT 36

US-09-810-997-8/C

Sequence 8, Application US/09810997

Patent No. US2002009501A1

GENERAL INFORMATION:

APPLICANT: Song, Xiaoling

APPLICANT: Fan, Hao

APPLICANT: Wei, Zhong-Min

TITLE OF INVENTION: PEPTIDES FOR HYPERSENSITIVE RESPONSE ELICITORS AND

FILE REFERENCE: 21829/62

CURRENT APPLICATION NUMBER: US/09/810,997

CURRENT FILING DATE: 2001-03-16

PRIOR APPLICATION NUMBER: 60/191,649

PRIOR FILING DATE: 2000-03-23

PRIOR APPLICATION NUMBER: 60/250,710

PRIOR FILING DATE: 2000-12-01

NUMBER OF SEQ ID NOS: 9

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 8

LENGTH: 342

TYPE: DNA

ORGANISM: Xanthomonas campestris pv. pelargonii

OTHER INFORMATION:

US-09-810-997-8

Alignment Scores:

Pred. No.: 113 Length: 342

Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.15% Indels: 0

DB: 10 Gaps: 0

US-09-831-458a-12 (1-325) x US-09-810-997-8 (1-342)

QY 32 LeuleuAlaGlyValLeuVal 38

DB 90 CTGCTGAGTGGAGTCTCTTCA 70

RESULT 37

US-09-880-371-14/C

Sequence 14, Application US/09880371

Patent No. US20020059658A1

GENERAL INFORMATION:

APPLICANT: Wei, Zhong-Min

APPLICANT: DeRoche, Jay

TITLE OF INVENTION: METHODS OF IMPROVING THE EFFECTIVENESS OF TRANSGENIC

FILE REFERENCE: 21829/91

CURRENT APPLICATION NUMBER: US/09/880,371

CURRENT FILING DATE: 2001-06-14

PRIOR APPLICATION NUMBER: 60/211,585

PRIOR FILING DATE: 2000-06-15

NUMBER OF SEQ ID NOS: 16

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 14

LENGTH: 342

TYPE: DNA

ORGANISM: Xanthomonas campestris pv. pelargonii

OTHER INFORMATION:

```

: TYPE: DNA
: ORGANISM: Xanthomonas campestris
US-09-880-371-14

Alignment Scores:
Pred. No.: 113
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.15%
DB: 10

US-09-831-458A-12 (1-325) x US-09-880-371-14 (1-342)

QY 32 LeuLeuAlaGlyValLeuVal 38
DB 90 CTGCTGCTGAGTCTGCTG 70

RESULT 38
US-09-880-371-15/C
: Sequence 15, Application US/09880371
: Patent No. US20020059658A1
: GENERAL INFORMATION:
: APPLICANT: Wei, Zhong-Min
: APPLICANT: Debocher, Jay
: TITLE OF INVENTION: METHODS OF IMPROVING THE EFFECTIVENESS OF TRANSGENIC
: FILE REFERENCE: 21829/91
: CURRENT APPLICATION NUMBER: US/09/880,371
: CURRENT FILING DATE: 2001-06-13
: PRIOR APPLICATION NUMBER: 60/211,585
: PRIOR FILING DATE: 2000-06-15
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: Patent Ver. 2.1
: SEQ ID NO 15
: LENGTH: 342
: TYPE: DNA
: ORGANISM: Nicotiana tabacum
US-09-880-371-15

Alignment Scores:
Pred. No.: 113
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.15%
DB: 10

US-09-831-458A-12 (1-325) x US-09-880-371-15 (1-342)

QY 32 LeuLeuAlaGlyValLeuVal 38
DB 90 CTGCTGCTGAGTCTGCTG 70

RESULT 39
US-09-829-124-1/C
: Sequence 1, Application US/09829124
: Patent No. US20020066122A1
: GENERAL INFORMATION:
: APPLICANT: Wei, Zhong-Min
: APPLICANT: Swanson, Shane S.
: TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR FROM XANTHOMAS
: FILE REFERENCE: 21829/101
: CURRENT APPLICATION NUMBER: US/09/829,124
: CURRENT FILING DATE: 2001-04-09
: PRIOR APPLICATION NUMBER: 60/224,053
: PRIOR FILING DATE: 2000-08-09
: PRIOR APPLICATION NUMBER: 09/412,452
: PRIOR FILING DATE: 1999-10-04
: PRIOR APPLICATION NUMBER: 60/103,124
: PRIOR FILING DATE: 1998-10-05

```

```

: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: Patent Ver. 2.1
: SEQ ID NO 1
: LENGTH: 342
: TYPE: DNA
: ORGANISM: Xanthomonas campestris
US-09-829-124-1

Alignment Scores:
Pred. No.: 113
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.15%
DB: 10

US-09-831-458A-12 (1-325) x US-09-829-124-1 (1-342)

QY 32 LeuLeuAlaGlyValLeuVal 38
DB 90 CTGCTGCTGAGTCTGCTG 70

RESULT 40
US-09-983-965-2185
: Sequence 2185, Application US/09983965
: Patent No. US20020137160A1
: GENERAL INFORMATION:
: APPLICANT: Warren, Wesley C.
: APPLICANT: Tao, Nengping
: APPLICANT: Byatt, John C.
: APPLICANT: Mathiadagu, Nandapan
: TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOC
: FILE REFERENCE: 37-21(110297)C
: CURRENT APPLICATION NUMBER: US/09/983,965
: CURRENT FILING DATE: 2001-10-26
: PRIOR APPLICATION NUMBER: US/09/465,241
: PRIOR FILING DATE: 1999-12-15
: PRIOR APPLICATION NUMBER: US/09/114,678
: PRIOR FILING DATE: 1998-12-17
: NUMBER OF SEQ ID NOS: 5912
: SEQ ID NO 2185
: LENGTH: 349
: TYPE: DNA
: ORGANISM: Bos taurus
: FEATURE:
: OTHER INFORMATION: Clone ID: 4 L18457-007-Q1-K1-04
US-09-983-965-2185

Alignment Scores:
Pred. No.: 115
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.15%
DB: 10

US-09-831-458A-12 (1-325) x US-09-983-965-2185 (1-349)

QY 24 ValLeuGlnLeuLeuSerPhe 30
DB 168 GTTTCAGCTGCTGAGTTT 168

RESULT 41
US-09-867-701-3187
: Sequence 3187, Application US/09867701
: Patent No. US2002013237A1
: GENERAL INFORMATION:
: APPLICANT: Aghate, Paul A.
: APPLICANT: Jones, Robert
: APPLICANT: Harlocker, Susan L.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: OTHER INFORMATION: AND DIAGNOSIS OF OVARIAN CANCER

```

```

: FILE REFERENCE: 210121.497
: CURRENT APPLICATION NUMBER: US/09/867,701
: CURRENT FILING DATE: 2001-05-29
: NUMBER OF SEQ ID NOS: 10912
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 3187
: LENGTH: 454
: TYPE: DNA
: ORGANISM: Homo sapien
US-09-867-701-3187

Alignment Scores:
Pred. No.: 117          Length: 354
Score: 7.00           Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 2.15%             Indels: 0
DB: 10                      Gaps: 0

US-09-831-458a-12 (1-325) x US-09-867-701-3187 (1-354)

QY 243 10cgccthrserarjserasu 249
DB 433 TCGCAAGACTCTAGAGCAAT 353

RESULT 42
US-09-880-107-1587/
: Sequence 158, Application US/09880107
: Patent No. US20020142981A1
: GENERAL INFORMATION:
: APPLICANT: Horne, Darci T.
: APPLICANT: Vockley, Joseph G.
: APPLICANT: Scherff, Uwe
: APPLICANT: Gene Logic, Inc.
: TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
: FILE REFERENCE: 44921.5028.WO
: CURRENT APPLICATION NUMBER: 09/09/880,107
: CURRENT FILING DATE: 2001-06-14
: PRIOR APPLICATION NUMBER: US 60/211,379
: PRIOR FILING DATE: 2000-06-14
: PRIOR APPLICATION NUMBER: US 60/237,054
: PRIOR FILING DATE: 2000-10-02
: NUMBER OF SEQ ID NOS: 3950
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 158
: LENGTH: 363
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: Genbank Accession No. US20020142981A1 AAC71387
US-09-880-107-158

Alignment Scores:
Pred. No.: 119          Length: 363
Score: 7.00           Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 2.15%             Indels: 0
DB: 10                      Gaps: 0

US-09-831-458a-12 (1-325) x US-09-880-107-158 (1-363)

QY 4 Setlysgluhrarqvalslu 10
DB 433 TCCAAAGAACCAAGACTGCCA 313

RESULT 43
US-09-878-574-182/
: Sequence 182, Application US/09878574
: Patent No. US20020110548A1
: GENERAL INFORMATION:
: APPLICANT: Byrum, Joseph R.
: APPLICANT: La Rosa, Thomas J.

: APPLICANT: Thompson, Michael D.
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
: FILE REFERENCE: 38-21(15401)B
: CURRENT APPLICATION NUMBER: US/09/878,574
: CURRENT FILING DATE: 2001-12-21
: PRIOR APPLICATION NUMBER: 09/434,535
: PRIOR FILING DATE: 1999-06-14
: NUMBER OF SEQ ID NOS: 15775
: SEQ ID NO 182
: LENGTH: 364
: TYPE: DNA
: ORGANISM: Glycine max
: OTHER INFORMATION: Clone ID: LIB3028-054-Q1-B1-D6
US-09-878-574-182

Alignment Scores:
Pred. No.: 120          Length: 364
Score: 7.00           Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 2.15%             Indels: 0
DB: 10                      Gaps: 0

US-09-831-458a-12 (1-325) x US-09-878-574-182 (1-364)

QY 271 scftrtscstgscptscscphc 277
DB 73 TCGCTACTTCACCATCTTC 53

RESULT 44
US-09-960-352-2949
: Sequence 2949, Application US/09960352
: Patent No. US20020147139A1
: GENERAL INFORMATION:
: APPLICANT: Warren, Wesley C.
: APPLICANT: Tao, Ningding
: APPLICANT: Byatt, John C.
: APPLICANT: Mathalaagan, Nagappan
: TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
: FILE REFERENCE: 16511.006/27-21(10298)C
: CURRENT APPLICATION NUMBER: 09/09/960,352
: CURRENT FILING DATE: 2001-09-24
: NUMBER OF SEQ ID NOS: 15112
: SEQ ID NO 2949
: LENGTH: 384
: TYPE: DNA
: ORGANISM: Bos taurus
: OTHER INFORMATION: Clone ID: 13 LTB14-607-01-E1-D1
US-09-960-352-2949

Alignment Scores:
Pred. No.: 126          Length: 484
Score: 7.00           Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 2.15%             Indels: 0
DB: 10                      Gaps: 0

US-09-831-458a-12 (1-325) x US-09-960-352-2949 (1-384)

QY 269 Aspclyscrtproleuscripto 275
DB 23 GATCGAACCCCTCTGCACCA 43

RESULT 45
US-09-826-124-477/
: Sequence 5, Application US/09826124
: Patent No. 002002006122A1
: GENERAL INFORMATION:
: APPLICANT: Wei, Zhong-Min
: APPLICANT: Swanson, Shane S.

```

APPLICANT: Fan, Hao
TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR FROM XANTHOMAS
FILE OF INVENTION: CAMPESTRIS
FILE REFERENCE: 21829/101
CURRENT APPLICATION NUMBER: US/09/829,124
CURRENT FILING DATE: 2001-04-09
PRIOR APPLICATION NUMBER: 60/224,053
PRIOR FILING DATE: 2000-08-09
PRIOR APPLICATION NUMBER: 09/412,452
PRIOR FILING DATE: 1999-10-04
PRIOR APPLICATION NUMBER: 60/103,124
PRIOR FILING DATE: 1998-10-05
NUMBER OF SEQ ID NOS: 6
SOFTWARE: patentIn Ver. 2.1
SEQ ID NO: 6
LENGTH: 408
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: probe
US-09-829-124-6

Alignment Scores:
Pred. No.: 134 Length: 408
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 10 Gaps: 0

US-09-831-458a-12 (1-325) X US-09-829-124-6 (1-408)

QY 32 LeuLeuAlaGlyValLeuVal 38
|||||
DB 142 CTGCTGGCTGGAGTCTCTGCTG 122

Search completed: December 7, 2002, 13:54:53
Job time : 90 secs

•
•
•
•

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd

OM protein - protein search, using SW model

Run on: December 7, 2002 11:00:15 Search time: 48 seconds
(without alignments)
1562184 Million cell updates/sec

Title: US-09-831-458A-12

Perfect score: 325

Sequence: 1 MSLSKPPRQQLNLGLSH PDDVNWYCKKPAACPRDE 325

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size: 5

Total number of hits satisfying chosen parameters: 17736

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Database: PIP-73.3

1: p1r1:
2: p1r2:
3: p1r3:
4: p1r4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	24.0	404	2 A46274	HIV gp120-binding
2	9	2.8	168	2 B88102	protein W09G10.5 (
3	8	2.5	214	2 E82350	orotate phosphorib
4	8	2.5	279	2 T52454	ATP-dependent Clp
5	8	2.5	282	2 S74644	regulatory compone
6	8	2.5	382	2 T01943	hypothetical prote
7	7	2.2	129	1 DEECS4	succinate dehydrog
8	7	2.2	129	1 H90722	succinate dehydrog
9	7	2.2	129	2 A85573	succinate dehydrog
10	7	2.2	139	2 AH0590	succinate dehydrog
11	7	2.2	130	2 A82119	succinate dehydrog
12	7	2.2	120	2 B70302	hypothetical prote
13	7	2.2	161	2 S76399	hypothetical prote
14	7	2.2	207	2 T21881	hypothetical prote
15	7	2.2	210	2 S43399	GTP-binding protei
16	7	2.2	227	2 T35281	probable two compo
17	7	2.2	232	2 D85654	probable urase ac
18	7	2.2	235	2 AF3451	1198 protein (limp
19	7	2.2	239	2 F71343	hypothetical prote
20	7	2.2	240	2 S65218	hypothetical prote
21	7	2.2	244	2 T26913	hypothetical prote
22	7	2.2	244	2 T26912	hypothetical prote
23	7	2.2	246	2 F95397	hypothetical prote
24	7	2.2	247	2 A99794	Probable halocid
25	7	2.2	247	2 G82956	urase associated
26	7	2.2	275	2 A64145	hypothetical prote
27	7	2.2	275	2 C81843	hypothetical prote
28	7	2.2	275	2 C81097	catboxylase (
29	7	2.2	275	2 A97226	esterase, probable
30	7	2.2	275	2 A97226	uncharacterized co

30	7	2.2	283	2 T04778	Protein W09G10.5 (
31	7	2.2	288	2 B97074	Protein W09G10.5 (
32	7	2.2	296	2 A22847	Protein W09G10.5 (
33	7	2.2	304	2 T20209	Protein W09G10.5 (
34	7	2.2	304	2 H81449	Protein W09G10.5 (
35	7	2.2	312	2 E57614	Protein W09G10.5 (
36	7	2.2	338	2 F71957	Protein W09G10.5 (
37	7	2.2	338	2 E64521	Protein W09G10.5 (
38	7	2.2	352	2 T05723	Protein W09G10.5 (
39	7	2.2	367	2 F64847	Protein W09G10.5 (
40	7	2.2	367	2 A85644	Protein W09G10.5 (
41	7	2.2	367	2 C90760	Protein W09G10.5 (
42	7	2.2	368	2 D64101	Protein W09G10.5 (
43	7	2.2	376	2 A95447	Protein W09G10.5 (
44	7	2.2	392	2 F96770	Protein W09G10.5 (
45	7	2.2	394	2 D86900	Protein W09G10.5 (

ALIGNMENTS

RESULT 1

A46274

HIV gp120-binding C-type lectin - human

C:Species: Homo sapiens (man)

C>Date: 21-Sep-1993 #sequence_statistics: 18 Nov 1994 #level: 100

C:Accession: A46274

R:Curis: B.M.; Scharnowski, S.; Watson, A.J.

Proc. Natl Acad. Sci. U S A 89, 8456-8460, 1992

A:Title: Sequence and expression of a membrane associated C-

A:Reference number: A46274; M010192304446; PMID:1518869

A:Accession: A46274

A>Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-404 <CUR>

A:Experimental source: placenta

A:Note: sequence extracted from NCBI backbone (NCBI:113134)

F:256-377/domain: C-type lectin homology - LCH.

Query Match

Best Local Similarity: 100.0%; Score: 78; DB 2; Length: 4

Matches: 78; Conservative: 0; Mismatches: 0; Indel:

QY 58 GATATGNTTCAKAAVGESEKSTAGHYELTGLKAAVGEPEKSKLQET

tb 74 GATATGNTTCAKAAVGESEKSKLQETGLKAAVGEPEKSKLQET

QY 118 VGELPEKSKLQETIVQELT 145

tb 134 VGELPEKSKLQETIVQELT 151

RESULT 2

B88102

protein W09G10.5 [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 10-May-2001 #sequence_statistics: 16 May 2001 #text: ch

C:Accession: B88102

C:Annotation: The C. elegans Sequencing Consortium.

A:Title: genome sequence of the nematode C. elegans: a plat

A:Reference number: A75080; M0101920613; PMID:9851916

A:Note: seq webstis genome webst. 10/29/97/2/eleans/ and we

A:Accession: B88102

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-168 <STO>

A:Cross-references: gtrchr_11: P00N:Abp013.1; PIP:q291566

C:Genetics:

A:Gene: W09G10.5

A:Map position: 2

A:Accession: A28836
 A:Molecule type: DNA
 A:Residues: 1-129 <MOO>
 A:Cross-references: GR X00980; NID:q42921, EITN:CAAC485.1, EID:q42924
 J. Gen. Microbiol. 132: 3239-3251, 1986
 A:Title: Transcript analysis of the citrate synthase and succinate dehydrogenase genes
 A:Reference number: 141112; MIMD:88009821, PMID:3309132
 A:Accession: 184546
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-129 <PES>
 A:Cross-references: GR M30989, NID:q42942; PTDN:AAA4546.1, EID:q495784
 R. Blatner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
 Science 277: 1453-1462, 1997
 A:Title: The complete genome sequence of *Escherichia coli* K-12
 A:Reference number: A64720; MIMD:97426617, PMID:9278503
 A:Accession: H64807
 A:Status: preliminary
 A:Molecule type: nucleic acid sequence not shown; translation not shown
 A:Residues: 1-129 <RLAT>
 A:Cross-references: GR AE050175, GR:000096; NID:q4786934; EIDN:AMC73815.1; PTD:q4786940;
 A:Experimental source: strain K-12, substrain MG1655
 A:Genetics:
 A:Gene: sbdc; cybA
 A:Map position: 17 min
 A:Complex: one of two hydrophobic anchor components of the succinate dehydrogenase comp
 A:Function:
 A:Pathway: tricarboxylic acid cycle
 A:Superfamily: succinate dehydrogenase 14k hydrophobic protein
 A:Keywords: heme; oxidoreductase; transmembrane protein; tricarboxylic acid cycle
 E:32.48/Domain: transmembrane #status predicted <TML>
 E:73.89/Domain: transmembrane #status predicted <TML>
 F:112-128/Domain: transmembrane #status predicted <TML>

Query Match 2 %; Score 7; DB 2; Length 129,
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 LLACVLY 38
 Db 122 LLACVLY 128

RESULT 8
 B90722
 succinate dehydrogenase [imported] - *Escherichia coli* (strain O157:H7, substrain RIMD 05
 C:Species: *Escherichia coli*
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 0; Aug-2001
 C:Accession: B90722
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 gasawara, N.; Yasunaga, T.; Kihara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 4: 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gene
 A:Reference number: A94629; MIMD:21156231; PMID:11258796
 A:Accession: B90722
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-129 <HAV>
 A:Cross-references: GR:BA000007; PTDN:BAH4169.1; PTD:q43760204; GSPR:GN00154
 A:Experimental source: strain O157:H7, substrain RIMD 0509452
 A:Genetics:
 A:Gene: Ecs0746
 C:Superfamily: succinate dehydrogenase 14k hydrophobic protein

Query Match 2 %; Score 7; DB 2; Length 129,
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 LLACVLY 38
 Db 122 LLACVLY 128

RESULT 9
 AB5573
 succinate dehydrogenase, cytochrome b556 [imported] - *Escherichia coli*
 C:Species: *Escherichia coli*
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 0
 C:Accession: AB5573
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasco
 111er, L.; Grobark, F.T.; Davis, N.W.; Lin, A.; Finkelstein, J.
 Nature 409: 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic *Escherichia coli*
 A:Reference number: AB5480; MIMD:11043457; PMID:11265551
 A:Accession: AB5573
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-129 <STO>
 A:Cross-references: GR AE050175; NID:q42942; PTDN:AA05504
 A:Experimental source: strain O157:H7, substrain EDL933
 A:Genetics:
 A:Gene: sbdc
 C:Superfamily: succinate dehydrogenase 14k hydrophobic protein

Query Match 2 %; Score 7; DB 2; Length 129,
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 LLACVLY 38
 Db 122 LLACVLY 128

RESULT 10
 AH0590
 succinate dehydrogenase cytochrome b 556 chain [imported] -
 C:Species: *Salmonella enterica* subsp. *enterica* serovar Typhimurium
 A:Note: this species has also been called *Salmonella typhi*
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 0
 C:Accession: AH0590
 R:Parkhill, J.; Bengtson, G.; James, K.D.; Thomson, N.P.; Pych
 ch, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, E.M.;
 S.; Mout, S.; O'Garra, P.
 Nature 413: 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.
 A:Title: Complete genome sequence of a multiple drug resistant
 A:Reference number: AB0502; PMID:11677608
 A:Accession: AH0590
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-129 <PAR>
 A:Cross-references: GR:AL513382; PTDN:CA005194.1; PTD:q416501
 A:Genetics:
 A:Gene: STY0775
 C:Superfamily: succinate dehydrogenase 14k hydrophobic protein

Query Match 2 %; Score 7; DB 2; Length 129,
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 LLACVLY 38
 Db 122 LLACVLY 128

RESULT 11
 AB2119
 succinate dehydrogenase, cytochrome b556 chain W2091 [imported]
 C:Species: *Vibrio cholerae*
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 0
 C:Accession: AB2119
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Gitt
 L, P.; Mekalanos, J.T.; Venter, L.C.; Fraser, C.M.

Nature 406: 427-483, 2000
 A>Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
 A:Reference number: A82035, MIM: 244678, PMID:11542301
 A:Accession: A82119
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-130 <HEI>
 A:Cross-references: GR AF004283; GR AF003852; NID:9606636; PDB:MAF95237.1; GSPDB:GN001
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VC2091
 A:Map position: 1
 C:Superfamily: succinate dehydrogenase 14K hydrophobic protein

Query Match 2.2% Score 7; DB 2; Length 130;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q# 32 LLAGVY 38
 |||||
 Db 123 LLAGVY 129

RESULT 12
 B70302
 Hypothetical protein aq_026 *Aquifex aeolicus*
 C:Species: *Aquifex aeolicus*
 C>Date: 08-May-1998 #sequence_revision: 08-May-1998 #text_change: 05-Nov-1999
 C:Accession: B70302
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lonox, A.L.; Graham, D.F.; O'N.
 Nature 392: 353-358, 1998
 A>Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.
 A:Reference number: A70300; MIM:98196666; PMID:9537320
 A:Accession: B70302
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-160 <AOE>
 A:Cross-references: GR AF000076; NID:9398779; PDB:MAV9410.1; PDB:9498286; GR AF00007
 A:Experimental source: strain VFS
 C:Genetics:
 A:Gene: aq_026

Query Match 2.2% Score 7; DB 2; Length 163;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q# 32 LLAGVY 38
 |||||
 Db 4 LLAGVY 10

RESULT 13
 S76399
 Hypothetical protein - *Synechocystis* sp. (strain PCC 6803)
 C:Species: *Synechocystis* sp.
 A:Variety: PCC 6803
 C>Date: 25-Apr-1997 #sequence_revision: 25-Apr-1997 #text_change: 0-Jun-2000
 C:Accession: S76399
 E:Kuroki, T.; Sato, S.; Kodani, H.; Tanaka, A.; Asanuma, E.; Nakamura, Y.; Miyajima, N.;
 O. K.; Okumura, S.; Shimpo, S.; Takuchi, G.; Wada, T.; Watanabe, A.; Yanada, M.; Yasuda
 DNA Res. 3: 109-136, 1996
 A>Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
 S.
 A:Reference number: S76322; MIM: 9706130; PMID:8905231
 A:Accession: S76399
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-161 <KAN>
 A:Address: Bethesda, MD: 20815, 58 AF00133; NID:9155054; NID:MAV9410.1; PDB:9498286
 A:Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C:Superfamily: *Synechocystis* Hypothetical protein slh0308

Query Match 2.2% Score 7; DB 2; Length 161;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q# 48 PSSISQ 54
 |||||
 Db 29 PSSISQ 45

RESULT 14
 T21881
 Hypothetical protein F36H1.5 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C>Date: 15-Oct-1999 #sequence_revision: 15-Oct-1999 #text_change: 15-Oct-1999
 C:Accession: T21881
 R:Korshaw, J.
 Submitted to the EMBL Data Library, January 1996
 A:Reference number: Z19482
 A:Accession: T21881
 A>Status: preliminary; translated from GR/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-207 <WIL>
 A:Cross-references: EMBL:Z68760; PDB:1CAV2993.1; GSPDB:GN00922; GSPDB:F36H1.5
 A:Experimental source: clone F36H1
 C:Genetics:
 A:Gene: CESP:F36H1.5
 A:Map position: 4
 A:Insertion: 124/1; 154/1

Query Match 2.2% Score 7; DB 2; Length 207;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q# 12 LGHGGI 18
 |||||
 Db 130 LGHGGI 136

RESULT 15
 S43399
 GTP-binding protein YPS21 - yeast (*Saccharomyces cerevisiae*)
 A:Alternate names: GTP-binding protein YPS15; protein YPS14; protein YPS089; protein
 C:Species: *Saccharomyces cerevisiae*
 C>Date: 11-Mar-1997 #sequence_revision: 14-Sep-1994 #text_change: 02-Feb-2001
 C:Accession: S43399; S4169; S4169; S61649; S66374; S40001
 E:Haradovsky, B.F.; Busch, G.R.; Emr, S.D.
 EMBO J. 13: 1297-1309, 1994
 A>Title: YPS21 encodes a Rab5-like GTP-binding protein that is required for the sorting
 A:Reference number: S41399; MIM:3418535; PMID:8137814
 A:Accession: S43399
 A:Molecule type: DNA
 A:Residues: 1-210 <HOR>
 A:Cross-references: EMBL:Z29438; NID:9441376; PDB:1CAV2993.1; PDB:9441477
 R:Shen, K.; R. Steinberg, H.; Dueschle, A.; Philippidis, P.; Yoo, J.S.; Gall
 J. Cell Biol. 125: 283-299, 1994
 A>Title: Role of three Rab5-like GTPases, YPS1p, YPS5p, and YPS3p, in the endocytic
 A:Reference number: S41399; MIM:3418535; PMID:8135546
 A:Accession: S43399
 A>Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-210 <STN>
 A:Cross-references: EMBL: X76173; NID:948366; PDB:1CAV2993.1; PDB:9483567
 E:Jones, V.; A. Garcia, M.A.; Rodriguez, S.; Rodicio, G.; Barreiro, A.; Sanchez, C.; Valen
 Submitted to the EMBL Data Library, December 1995
 A:Description: Nucleotide sequence and analysis of a 130 kb fragment of yeast chromosome
 A:Reference number: S61649
 A:Accession: S61649
 A:Molecule type: DNA
 A:Residues: 1-210 <REN>
 A:Cross-references: EMBL: X64457; NID:9126179; PDB:1CAV2993.1; PDB:91164935
 E:Yoss, H.; Jones, V.; Rodicio, S.; Rodicio, G.; Schwagerl, C.; Paves, V.; Ausorge, W
 Submitted to the Protein Sequence Database, July 1996
 A:Reference number: S66965

A:Accession: S66974
 A:Molecule type: DNA
 A:Cross-references: GR:AE005174, NID:q12514461, PIRN:AA055287
 A:Experimental source: strain S288C
 C:Genetics:
 A:Gene: SCD:VPS21: YPT51
 A:Cross-references: SGD:S0005615; MIPS:Y0P089c
 A:Map position: 15R
 C:Superfamily: ras transforming protein; translation elongation factor Tu homology
 C:Keywords: GTP binding; membrane trafficking; nucleotide binding; P-loop
 F:14-21/Region: nucleotide-binding motif A (P-loop)
 F:120-123/Region: GTP-binding NKXD motif
 F:151-155/Region: GTP-binding SAK/L motif
 F:20/Binding site: GTP (Lys) #status predicted

Query Match 2.2% Score 7; DB 2; Length 210;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 7; Conservative 0; Mismatches 0; Indels 0 Gaps 0;

OY 233 KTAEEUN 239
 |||||
 Db 174 KTAEEUN 180

RESULT 16
 Probable two component response regulator - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #extl_change 31-Jan-2000
 C:Accession: T35281
 R:Seeger, K.; Harris, D.; Bentley, S.P.; Parkhill, J.; Barrett, R.; Rajadream, M.A.
 Submitted to the EMBL data library, July 1999
 A:Reference number: Z21574
 A:Accession: T35281
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-227 <SEE>
 A:Cross-references: EMBL:AL096872; PIRN:CA051249.1; GSPR:GN00070; NID:q12514461; MIPS:Y0P089c
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCD:SRB:SCSP7.36c
 C:Superfamily: regulatory protein ccmA, response regulator homology

Query Match 2.2% Score 7; DB 2; Length 227;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 7; Conservative 0; Mismatches 0; Indels 0 Gaps 0;

OY 226 RAQIV 242
 |||||
 Db 209 RAQIV 215

RESULT 17
 Probable urease accessory protein B (imported) - Escherichia coli (strain 0157:H7, subst
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #extl_change 14-Sep-2001
 C:Accession: D85654
 R:Perma, N.T.; Plunkett III, G.; Rutland, V.; Mau, R.; Gysner, J.F.; Pese, P.T.; Mayhew
 Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimarata, E.; Petamoussis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli 0157:H7
 A:Reference number: AB5480; MIMD:21074935; PMID:11206551
 A:Accession: D85654
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-232 <STO>
 A:Cross-references: GB:AE005174, NID:q12514461, PIRN:AA055287.1, GSPR:GN00145; MWCP:215
 A:Experimental source: strain 0157:H7, substrain EDL933
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-232 <STO>
 A:Cross-references: GR:AE005174, NID:q12514461, PIRN:AA055287
 A:Experimental source: strain 0157:H7, substrain EDL933
 C:Genetics:
 A:Gene: ureD_2; ureD

Query Match 2.2% Score 7; DB 2; Length 24;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 7; Conservative 0; Mismatches 0; Indels 0 Gaps 0;

OY 267 WVDGSP 273
 |||||
 Db 156 WVDGSP 162

RESULT 18
 IrgB protease (imported) - Brucella melitensis (strain 16M)
 C:Species: Brucella melitensis
 C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #extl_change 1
 C:Accession: AF3451
 R:DeIvecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Ma
 Mazur, M.; Goldsman, E.; Selkov, E.; Elver, P.H.; Hagius,
 Proc. Natl. Acad. Sci. U.S.A. 99, 4444-4448, 2002
 A:Title: The genome sequence of the facultative intracellular
 A:Reference number: AD1252; PMID:11766688
 A:Accession: AF3451
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-235 <KRS>
 A:Cross-references: GR:AE008917; PIRN:AA152777.1; PIR:q17980
 A:Experimental source: strain 16M
 C:Genetics:
 A:Gene: BME11596
 A:Map position: 1
 C:Superfamily: yolk protein

Query Match 2.2% Score 7; DB 2; Length 24;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 7; Conservative 0; Mismatches 0; Indels 0 Gaps 0;

OY 32 LAGIV 38
 |||||
 Db 100 LAGIV 106

RESULT 19
 E7143
 Hypothetical protein H7282 - syphilis spirochete
 C:Species: Treponema pallidum subsp. pallidum (syphilis sp)
 C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #extl_change 1
 C:Accession: E7143
 R:Fraser, C.M.; Norris, S.J.; Welstock, G.M.; White, O.; Se
 rson, J.; Khaliq, H.; Richardson, D.; Howell, J.K.; Chidamb
 they, L.; Weisman, J.; Smith, H.O.; Wenter, J.C.
 Science 281, 375-388, 1998
 A:Title: Complete genome sequence of Treponema pallidum, the
 A:Reference number: AF1250; MIMD:98442770; PMID:9665876
 A:Accession: E7143
 A:Status: preliminary; nucleic acid sequence not shown; trans
 A:Molecule type: DNA
 A:Residues: 1-239 <COL>
 A:Cross-references: GR:AF001209; GR:AE005203; NID:q12514461
 A:Experimental source: strain Nichols
 C:Genetics:
 A:Gene: TP0282
 C:Superfamily: syphilis spirochete hypothetical protein (pp)

Query Match 2.2% Score 7; DB 2; Length 24;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 7; Conservative 0; Mismatches 0; Indels 0 Gaps 0;

OY 35 GIVAIL 41


```

A:Reference number: A96000; MW:21350.325; PMD:21359.325
A:Accession: A97226
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-275 <KOR>
A:Cross-References: GR-AE001437; F1FN-AAE8566_1; F1D-915025477; SAEH-55N-0148
A:Experimental source: Clostridium acetobutylicum ATCC824
A:Genetics:
A:Gene: CAC2649

Query Match
Best Local Similarity 100.0%; Score 7; DB 2; Length 275;
Matches 7; Conservative 0; Mismatches 0; Gaps 0;

QY 22 AAAYQLL 28
DB 92 AAVGSH 98

RESULT 30
105778
peroxidase (EC 1.11.1.7), seed coat - soybean (fragment)
C:Species: Glycine max (soybean)
C:Date: 23 Apr 1999 #sequence_revision 23 Apr 1999 #text_change 03 Mar 2000
C:Accession: T06778
R:Inanpu, J.; Graham, M.C.; Graham, J.S.
submitted to the EMBL Data Library, November 1995
A:Reference number: Z15805
A:Accession: T06778
A>Status: preliminary; translated from GB/EMBL/DD8J
A:Molecule type: mRNA
A:Residues: 1-283 <HMA>
A:Cross-References: EMBL:J04167; M1D-9157103; F1FN-AAH9129_1; F1D-9155484
A:Experimental source: cultivar Williams 280uhghly
A:Genetics:
A:Note: SP0D4.1
A:Description: catalyzes the oxidation of a donor using hydroperoxide as electron ac
C:Superfamily: peroxidase
C:Keywords: heme, iron, metalloprotein, oxidoreductase
F12.129/Binding site: heme iron (HIS) (axial ligands) #status predicted
F14.5/Disulfide bonds: #status predicted
F157-256/Disulfide bonds: #status predicted
F16-166/Disulfide bonds: #status predicted

Query Match
Best Local Similarity 100.0%; Score 7; DB 2; Length 283;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 NITQAKA 70
DB 104 NITQAKA 110

RESULT 31
H97074
uncharacterized proteins, homologs of microcin C7 resistance protein MccF (F09414)
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14 Sep 2001 #text_change 14 Sep 2001
C:Accession: H97074
R:Rolling, J.; Berton, G.; Omelchenko, M.V.; Markarova, R.S.; Zedler, J.; Gibson, B.; Lee,
J.; Daly, M.T.; Bennett, G.N.; Koonin, E.V.; Smith, D.P.
J. Bacteriol. 183, 4823-4838, 2001
A>Title: Genome Sequence and Comparative Analysis of the Solvent Producing Bacterium Clo
A:Feature number: A96000; MW:21350.325; PMD:21359.325
A:Accession: H97074
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-288 <KUP>
A:Cross-References: GB-AE001437; F1FN-AAH74687_1; F1D-915524577; SAEH-55N-0148
A:Experimental source: Clostridium acetobutylicum ATCC824
A:Genetics:
A:Gene: CAC1419

```

```

Query Match:          2.2%   Score 76   EB 2:   Length 288;
Best Local Similarity 100.0%; Pred. No. 56;
Matches      7; Conservative    0; Mismatches     0; Indels      0; Gaps      0;

CY       295 AEFSSGSG 401
        |||
Db        54 AEFSSGSG 60

RESULT 32
AC2837
C:10-methyltransferase [unreviewed] - Agrobacterium tumefaciens (strain C)
C:Species: Agrobacterium tumefaciens
C:Dated: Mar-2002 #sequence_provision 11 Jan-2002 #text_change 0; Feb-2002
C:Accession: AC2837
E:Wood, J.W., Schmitt, J.C., Kaul, R., Marks, D., Chen, L., Mendel, G.F., Chen, Y., Woo
erger, G.J., Gillet, W., Grant, C.J., Guenther, D., Kutayavin, T., Levy, R., Li, M., McCl
rarp, P., Komatsu, P., Zhang, S.
Science 294, 237-243, 2001
A:Authors: Wood, J.W., Tao, Y., Biddle, P., Jung, M., Krespan, W., Perry, M., Gordon-Kam
ster, P.W.
A>Title: The genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AF0677, PMID:174193
A:Accession: AC2837
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-296 cKbps
A:Cross-reference: ChAF068688, EMBL AA141133.1, F01-g17740585; GS008-GN00166
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: melf
A:Map location: Chromosome
C:Superfamily: 5.19 methyltransferase and reductase (FADH2)

Query Match:          2.2%   Score 76   DB 2:   Length 286;
Best Local Similarity 100.0%; Pred. No. 57;
Matches      7; Conservative    0; Mismatches     0; Indels      0; Gaps      0;

CY       63 CNILGLE 69
        |||
Db        211 CNILGLK 217

RESULT 33
JX0209
C:Lectin, galactose-binding agglutinin from house mouse
C:Species: Mus musculus (house mouse)
C:Date: 20 Jun 1992 #sequence_provision 30-Mar-1992 #text_change 20-Aug-1999
C:Accession: JX0209; PX0009
A:Saito, M.; Kawakami, K.; Osawa, I.; Toyoshima, S.
J. Biochem. 111, 331-336, 1992
A:Reference number: JX0209; MIMD-ep09876; PMID:1587704
A:Accession: JX0209
A:Molecule type: mRNA
A:Status: 1-394 cSAR>
A:Cross-reference: JX0209; MIMD-ep09876; EMBL AB022771.1; F01-g249361
A:Title: Purification and characterization of a lectin-like molecule specific for gal
toida, S.; Saito, M.; Toyoshima, S.; Osawa, I.
J. Biochem. 104, 600-605, 1988
A:Reference number: FX0009; MIMD-ep09876; PMID:3241002
A:Accession: PX0009
A:Molecule type: protein
A:Status: 102-120-137-X; 139-171 cDA
A:Superfamily: hepatic lectin c-type lectin homology
C:Keywords: glycoprotein; lectin; macrophage; transmembrane protein
E:F36/61/2 domain; transmembrane *status predicted cTAS>
E:F173-226/blood; c-type lectin homology cLCH>
E:174-106/blood; c-type lectin homology (Asu) (cvalent) *status predicted

Query Match:          2.2%   Score 76   DB 2:   Length 304;
Best Local Similarity 100.0%; Pred. No. 58;

```


R. Barnett, M.J.; Fisher, R.F.; Jones, T.; Kemp, C.; Abola, A.P.; Barlow-Hubler, F.; Pows,
K.; Kaim, S.; Keating, D.H.; Palm, C.; Pock, M.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A>Title: Nucleotide sequence and predicted functions of the entire *Shorhizobium meliloti*
A:Reference number: A95262; MUID:2136509; PMID:1148132
A:Accession: A95147
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-376 <STO>
A:Cross-references: GB:AE006496; FIDN:AKR6333; FID:91453792; 3446-000395
A:Experimental source: Strain 1021, mesoplasmaid F57MA
A:Galat, F.; Flann, T.M.; Long, S.P.; Puhler, A.; Abola, P.; Amco, F.; Barlow-Hubler,
F.; Chait, P.; Cowie, A.; Davis, E.W.; Duncan, S.; Federspiel, N.A.; Fisher, R.F.;
Hyman, R.W.; Jones, T.;
Science 293, 908-922, 2001
A:Author: Kahn, D.; Kahn, M.; Kaim, S.; Keating, D.H.; Kiss, E.; Kemp, C.; Leclaire,
Hebault, P.; Vandenbol, M.; Vohlander, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A>Title: The complete genome of the legume symbiont *Shorhizobium meliloti*.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: nitr
A:Genome: plasmid
C:Keywords: oxidoreductase

Query Match 2.28; Score 7; DB 2; Length 376,
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 147 LKAAVGE 143
|||||
DB 209 LKAAVGE 275

RESULT 44

D96770
Hypothetical protein F1017.12 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D96770
R:Tipolopis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, C.; Alonso,
Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ausen, N.F.; Hughes, B.; Huizart, L.
Nature 408, 816-820, 2000
A:Author: Hunter, T.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Linds, T.S.; Maiti, P.; Matzall,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Author: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96770
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-592 <STO>
A:Cross-references: GB:AE005173; NID:96949233; FIDN:AAF1735.1; GSPDB:GN00141
C:Genetics:
A:Gene: F1017.12
A:Map position: 1

Query Match 2.28; Score 7; DB 2; Length 392;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 241 LQLOTSR 247
|||||
DB 37 LQLOTSR 43

RESULT 45

D86900
ABC Transporter permease protein ywH [imported] - Lactococcus lactis subsp. lactis (str
C:Species: Lactococcus lactis subsp. lactis

C:Date: 28-Mar-2001 #sequence_revision 28-Mar-2001 #text_change 28-Mar-2001
C:Accession: D86900
R:Bohlin, A.; Winkler, P.; Mauger, S.; Jallou, O.; Malarne,
Genome Res. 11, 731-753, 2001
A>Title: The complete genome sequence of the lactic acid bacterium
A:Reference number: A86625; MUID:21245186; PMID:11437471
A:Accession: D86900
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-394 <STO>
A:Cross-references: GB:AE005173; FID:91453792; FIDN:AKR6333;
A:Experimental source: Strain 111404
C:Genetics:
A:Gene: ywH

Query Match 2.28; Score 7; DB 2; Length 392,
Best Local Similarity 100.0%; Pred. No. 74;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 256 LSDNQE 262
|||||
DB 88 LSDNQE 94

Search completed: December 7, 2002, 11:17:42
Job time : 30 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2002, 11:07:40, Search time 17 seconds

(without alignments)
792,930 Million cell updates/sec

Title: US-09-831-458A-12

Sequence: 1 MSDSKPPVQGLGLGGLGH PPTVNNYWGKPPACFIDE 325

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Search: 112892 seqs, 41476328 residues

Word size: 5

Total number of hits satisfying chosen parameters: 7227

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	956	2.8	1	AD19_HUMAN	Q9H013 Homo sapien
2	214	2.5	1	FYRE_VIBCH	Q9KVD5 Vibrio chol
3	129	2.2	1	DHSC_ECOLI	P10446 Escherichia
4	129	2.2	1	DHSC_SALTY	O8X440 Salmonella
5	179	2.2	1	OPF3_HUMAN	Q9H6K4 Homo sapien
6	210	2.2	1	YF51_YEAST	P36017 Saccharomyc
7	239	2.2	1	Y283_TREPA	O83306 Leptotheca
8	275	2.2	1	YAIM_HAEIN	P44356 Haemophilus
9	304	2.2	1	MMGL_MOUSE	P49300 Mus musculu
10	367	2.2	1	PROB_PASMU	O9C105 Pasteurella
11	368	2.2	1	PROB_HAEIN	P43763 Haemophilus
12	390	2.2	1	YCCW_FUSNN	O81622 Fusobacteri
13	396	2.2	1	YCCW_ECOLI	P75676 Escherichia
14	471	2.2	1	PTMB_BACST	P50852 Bacillus st
15	497	2.2	1	DHAL_ASPNG	P41751 Aspergillus
16	501	2.2	1	CSO2_DROME	Q9VMT6 Drosophila
17	505	2.2	1	TCMO_POPKI	O43054 Populus kit
18	527	2.2	1	IMA_VYCES	O23478 Lycopersic
19	548	2.2	1	LACI_PHLKA	O01579 Phlebia rad
20	583	2.2	1	FR12_GROVI	O24760 Drosophila
21	610	2.2	1	PTMA_BACST	P42956 Bacillus su
22	708	2.2	1	ABBI_MOUSE	O94911 Mus musculu
23	710	2.2	1	ABBI_HUMAN	O00213 Mus sapien
24	715	2.2	1	ERF2_CANAL	O13354 Candida alb
25	720	2.2	1	TCME_HUMAN	O43548 Homo sapien
26	759	2.2	1	ITB3_BOVIN	P32552 Bos taurus
27	769	2.2	1	ITB3_HUMAN	P05107 Bos sapien
28	769	2.2	1	ITB3_PIG	P53714 Sus scrofa
29	798	2.2	1	ITB3_HUMAN	P26010 Homo sapien
30	804	2.2	1	SCV1_YEAST	P53009 Saccharomyc
31	804	2.2	1	YBHF_ECOLI	P77504 Escherichia
32	806	2.2	1	ITB7_MOUSE	P26011 Mus musculu
33	991	2.2	1	CLK1_DROME	Q03445 Drosophila

34	7	2.2	3414	1	PEL4_TREBH
35	6	1.8	84	1	BECL1_HUMAN
36	6	1.8	104	1	KK21_PORPU
37	6	1.8	123	1	Q59A_MOUSE
38	6	1.8	124	1	R121_SINY3
39	6	1.8	125	1	AWA3_TRLAB
40	6	1.8	129	1	Q59B_MOUSE
41	6	1.8	129	1	CYGD_RHOGE
42	6	1.8	132	1	R17A_PIG
43	6	1.8	140	1	ZG49_XENLA
44	6	1.8	141	1	HRM1_PORBA
45	6	1.8	141	1	HBP1_DRECA

ALIGNMENTS

RESULT 1
AD19_HUMAN
ID AD19_HUMAN STANDARD: PRT: 956 AA.
AC Q9H013; Q9H0P2; Q9H2L5;
DT 15-JUN-2002 (Ref. 41, Created)
DT 15-JUN-2002 (Ref. 41, Last sequence update)
DT 15-JUN-2002 (Ref. 41, Last annotation update)
DE ADAM 19 precursor (EC 3.4.24.-) (A disintegrin and meta-
domain 19) (Mellitin beta) (Metalloprotease and disintegrin
antigen marker) (MADAM).
DE ADAM19 OR MTM9 OR FKSG34.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eui-
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Hui-
NX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM A).
RA Wang Y.-G., Gong L.;
RT "Identification of FKSG34, a novel human gene encoding a
metalloprotease-disintegrin mellitin beta."
RL Submitted (DEC-2000) to the FMB/Genbank/EMBL databases
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM B).
RC TISSUE=Lymph node;
RX MEDLINE=20346928; Pubmed 1087142;
RA Fritzsche J., Moser M., Faust S., Becker A., Huettner R.,
Kreutz M.;
RT "Molecular cloning and characterization of a human meta-
disintegrin a novel marker for dendritic cell different.
KL Hlood 96:742-749(2000).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM B).
RC TISSUE=Dendritic cell;
RX MEDLINE=21094672; Pubmed 11125644;
RA Wei P., Zhao Y.-G., Zhuang L., Kolon S., Sang Q.-X.A.;
RT "Expression and enzymatic activity of human disintegrin
K1 metalloproteinase ADAM19/mellitin beta."
PL Biochem. Biophys. Res. Commun. 290:744-755(2001).
RN [4]
RP SEQUENCE OF 100-956 FROM N.A. (ISOFORM A).
RA Xu R., Cai J., Ying B., Wang F., Xu T., Zhao S., Li C.;
RT "Partial sequence of Homo sapiens ADAM19."
KL Submitted (MAR-1999) to the FMB/Genbank/EMBL databases
CC FUNCTION: PARTICIPATES IN THE PROTEOLYTIC PROCESSING
CC NEURULIN ISOFORMS WHICH ARE INVOLVED IN NEUROGENE-
CC SNAPISGENESIS, SUGGESTING A REGULATORY ROLE IN GLIA
CC CLEAVES ALPHA-2 MACROGLOBULIN. MAY BE INVOLVED IN
CC DIFFERENTIATION AND/OR OSTEOBLAST ACTIVITY IN BONE
CC (SIMILARITY).
CC -1- COFACTOR: BINDS ONE 21N-ION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type 1 membrane protein.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms: A (shown here) and
CC -1- TISSUE SPECIFICITY: EXPRESSED IN MANY NORMAL ORGANS
CC SEVERAL CANCER CELL LINES.
CC -1- INDUCTION: BY 1,25(OH)2VIT3 IN MONOCYTES.

```

01 -1- THE PRECURSOR IS CLEAVED BY A FUFIN ENDOPEPTIDASE (BY
02 SIMILARITY).
03 -1- SIMILARITY. BELONGS TO PEPTIDASE FAMILY M12B.
04 -1- SIMILARITY. CONTAINS 1 EGF-LIKE DOMAIN
05 -1- SIMILARITY. CONTAINS 1 EGF-REGRIN DOMAIN
06 -----
07 This SWISS-PROT entry is copyright. It is produced through a collaboration
08 between the Swiss Institute of Bioinformatics and the EMBL outstation in
09 the European Bioinformatics Institute. There are no restrictions on its
10 use by non-profit institutions as long as its content is in no way
11 modified and this statement is not removed. Usage by and for commercial
12 entities requires a license agreement (see http://www.isb-sib.ch/announce/
13 or send an email to license@isb-sib.ch).
14 -----
15 EMBL: AF326918; AAC50282.1; -
16 EMBL: Y13786; CAC20585.1; -
17 EMBL: AF131317; AAK07852.1; -
18 EMBL: AF134707; AAF22162.1; -
19 HSSP: P18619; 1EVL.
20 Genew: BCNC:197; ADAMI9
21 MIM: 603640; -
22 Interpro: IPR001762; Disintegrin
23 Interpro: IPR000561; EGF-like
24 Interpro: IPR001818; Matrixin
25 Interpro: IPR002870; Ref-M12b-Propep.
26 Interpro: IPR001500; Repolyisin
27 Interpro: IPR001190; Zn_MTPProtease.
28 Pfam: PF00200; disintegrin_1.
29 Pfam: PF01421; Repolyisin_1.
30 Pfam: PF15523; Pcp_M12b_Propep_1
31 PRINTS: PR00289; Disintegrin
32 Prodom: PD000664; Disintegrin_1.
33 SMART: SM00050; DISIN_1.
34 PROSITE: PS00215; ADAM_MPRO_1.
35 PROSITE: PS00546; CYSTEINE_SWITCH; FALSE_NEG
36 PROSITE: PS00427; DISINTEGRIN_1; FALSE_NEG.
37 PROSITE: PS00214; DISINTEGRIN_2; 1
38 PROSITE: PS00022; EGF_1; FALSE_NEG.
39 PROSITE: PS01186; EGF_2; 1.
40 PROSITE: PS00142; ZINC_PROTEASE; 1.
41 Hydroxase_Metalloprotease_E100_Sign_1; Glycophorin_3b_Membrane
42 Transmembrane; EGF-like domain; SH3-binding; Alternative splicing.
43 ET SIGNAL. 1 25
44 ET PROPEP. 26 203
45 ET CHAIN. 204 956
46 ET DOMAIN. 204 700
47 TRANSMEM. 701 721
48 ET DOMAIN. 722 956
49 ET DOMAIN. 204 410
50 ET DOMAIN. 416 502
51 ET DOMAIN. 435 438
52 ET DOMAIN. 503 660
53 ET DOMAIN. 651 663
54 ET SITE. 834 840
55 ET SITE. 839 845
56 ET SITE. 133 133
57 ET SITE. 346 346
58 ET ACT_SITE. 347 347
59 ET METAL. 350 350
60 ET METAL. 356 356
61 ET DISULFID. 321 404
62 ET DISULFID. 361 388
63 ET DISULFID. 475 482
64 ET DISULFID. 655 665
65 ET DISULFID. 659 671
66 ET DISULFID. 673 682
67 ET CARBOHYD. 145 145
68 ET CARBOHYD. 445 445
69 ET CARBOHYD. 448 448
70 ET CARBOHYD. 646 646
71 VASNYLIC. 903 956
72 -----
73 KPAKHSCTPLPA -> FPEYSGRAMTMSRT (IN
74 ISOFORM B).

```

```

FT  CONFLICT 4 4 C->S (IN REF. 2)
ET  CONFLICT 32 33 SK->R (IN REF. 2 AND 3)
ET  CONFLICT 558 558 V->D (IN REF. 2 AND 3)
ET  CONFLICT 623 623 N->D (IN REF. 2 AND 3)
SQ  SEQUENCE 956 AA; 105038 MW; 873810FA0418B12 CRC64;

Query Match: 7.8%; Score 9; DB 1; Length 956;
Best Local Similarity 100.0%; Pred No 0.79;
Matches 9, Conservative 6, Mismatches 0, Indels 0, Gaps 0;
QY 34 AGYVALVATV 42
    |||||
DB 706 AGYVALVATV 714

```

RESULT 2	SEQUENCE	SCORE	ALIGNMENT	IDENTITY	POSITIVES	NEGATIVES	MISSING	REMARKS
AC	Q9K9V95	214	100	100	0	0	0	100%
DT	15-JUN-2002 (Rel. 41, Created)							
DT	15-JUN-2002 (Rel. 41, Last sequence update)							
DE	Orotate Phosphoribosyltransferase (EC 2.4.2.16) (OPRT) (OPRTase).							
OS	Vibrio cholerae.							
CC	bacterial, Proteobacteria, gamma subdivision, Vibrionaceae: Vibrio.							
OX	NCBI_taxid:666;							
RN	[1]							
RP	SEQUENCE FROM N.A.							
RL	STRAIN: EL 101 N1691 / serotype O1;							
RX	MEDLINE: 20406833; Pubmed 10752301;							
FA	Heldreich J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gilm M.L.,							
RA	Indson R.J., Hart P.H., Huxley F.K., Peterson J.H., Umayam L.A.,							
RA	Gill S.R., Nishigaki K., Bond J.D., Heldreich J.F., Peterson J.H.,							
PA	Prinopaldeva M.D., Vamaachan J., Pass S., Qin H., Prasad I., Sellers P.,							
PA	McDonald J.D., Hitephack T., Fleischmann R.D., Newman W.C., White O.,							
PA	Salberg S.L., Smith H.C., Colwell P.P., Mekalanos J.T., Venturi T.C.,							
RA	Elaser C.M.;							
RT	"DNA sequence of both chromosomes of the cholera pathogen Vibrio							
RT	cholerae."							
AL	Nature 404:77-84(2000).							
CC	-1- CATALYTIC ACTIVITY: orotate 5'-phosphate + diphosphate -> orotate							
CC	+ 5-phospho-alpha-D-ribose 1-diphosphate.							
CC	! ALFABET: Lysimulase biosynthesis, fifth step.							
CC	! SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE							
CC	PHOSPHORIBOSYLTRANSFERASE FAMILY.							
CC								
CC	This Swiss-Prot entry is complete. It is produced through a collaboration							
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation in							
CC	the European Bioinformatics Institute. There are no restrictions on its							
CC	use by non-profit institutions as long as its content is in no way							
CC	modified and this statement is not removed. Usage by and for commercial							
CC	entities requires a license agreement (see http://www.isb.slb.org/submit/							
CC	or send an email to license@isb.slb.ch)							
DR	EMBL: AEO0411; AAF93387.1;							
DR	HSSP: P00495; I080.							
DR	TIGR: VC0211;							
DR	InterPro: IPR02375; Pf/Pf-IP-transf.							
DR	InterPro: IPR000836; PRtransferase.							
DR	InterPro: IPR004457; Pyre.							
DR	PIfam: PF00156; PRibosyltransf.1.							
DR	TIGRFAMs: TIGR00336; PYRE.1.							
DR	PROSITE: PS01093; P00_PYR_IP_TRANSFEP.1.							
KW	pyrimidine biosynthesis, transferase, glycosyltransferase;							
KW	Complete proteome.							
FT	ACT_SITE 104 .. 104							
FT	SEQUENCE 214 AA; 25544 MW; 6187171xColPrbAAV c70664;							
FT	BY SIMILARITY.							
QU	Query Match 2.5%; Score 8; DR 1; Length 214;							
QU	Best Local Similarity 100.0%; Pos. No. 2;							
QU	Mismatches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;							

```

QY      44 LAGVLVAI 40
DB      148 LAGVLVAI 155

RESULT 4
DHSC_EC01 STANDARD: PRT: 129 AA
AC P10446:
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE Succinate dehydrogenase cytochrome b-556 subunit.
GN SMC OR CYBA OR B0721 OR Z0875 OR EC00746.
OS Escherichia coli, and
OC Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83374;
RN
RP SEQUENCE FROM N.A.
RC STRAIN-K12:
RX MEDLINE=84307466; PubMed=6443359;
RA Wood D., Darlison M.G., Wilde R.J., Guest J.R.,
RT "Nucleotide sequence encoding the flavoprotein and hydrophobic
RL subunits of the succinate dehydrogenase of Escherichia coli";
RN Blochem. J. 222:519-534(1984).
RN
RP SEQUENCE FROM N.A.
RC STRAIN-K12:
RX MEDLINE=8809821; PubMed=3309132;
RA Wilde R.J., Guest J.R.;
RT "Transcript analysis of the citrate synthase and succinate
RL dehydrogenase genes of Escherichia coli K12.";
RN J Gen Microbiol. 132:3239-3251(1986).
RN
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655:
RX MEDLINE=97426617; PubMed=9278503;
RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Peto C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose B.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RN Science 277:1453-1474(1997).
RN
RP SEQUENCE FROM N.A.
RC STRAIN-K12:
RX MEDLINE=97061202; PubMed 8905232.
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mochizuki K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sanei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 7.18-kb DNA sequence of the Escherichia coli K-12 genome
RL corresponding to the 12.7-28.0 min region on the linkage map.";
RN DNA Res. 3:137-155(1996).
RN
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / FDI933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postel G., Hayckell J., Klink S., Boulton A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamasis K.,
RA Apodaca I., Anantharaman T.S., Lin J., Yee G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RN Nature 409:529-533(2001).

```

```

RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M.
RA Iida T., Takami H., Honda T., Sasaki G., Ogasawara N.,
RA Kohara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia
RL O157:H7 and genomic comparison with a laboratory strain";
RN DNA Res. 8:111-22(2001).
RN
RP IDENTIFICATION OF PROTEIN.
RA Murakami H., Kika K., Oya H., Aizaki Y.;
RT "The Escherichia coli cytochrome b556 gene, cybA, is ass.
RL in the succinate dehydrogenase gene cluster.";
RN FEMS Microbiol. Lett. 30:307-311(1985).
RN
RP MUTAGENESIS OF HISTIDINE RESIDUES.
RX MEDLINE=98189991; PubMed=9521746;
RA Vibat C.P., Greenhill G., Napatong K., Kita K., Dennis R.;
RT "Localization of histidine residues responsible for hemo-
RL ligation in cytochrome b556 of complex II (succinate dehy-
RL oxidoreductase) in Escherichia coli.";
RL Biochemistry 37:4148-4159(1998).
CC
CC FUNCTION: MONO-HEME CYTOCHROME OF THE SUCCINATE DEHY-
CC COMPLEX.
CC
CC PATHWAY: Titratoxyllic acid cycle.
CC
CC SUBUNIT: PART OF AN ENZYME COMPLEX CONTAINING FOUR S-
CC FLAVO-PROTEIN, AN IRON-SULFUR, CYTOCHROME B-556, AND AN
CC ANCHOR PROTEIN.
CC
CC SIMILARITY: BELONGS TO THE CYTOCHROME B556 FAMILY.
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through
CC between the Swiss Institute of Bioinformatics and the
CC the European Bioinformatics Institute. There are no re-
CC use by non-profit institutions as long as its content
CC modified and this statement is not removed. Usage by
CC entities requires a license agreement (See http://www.bf-
CC or send an email to license@isb.ch).
CC
CC
CC EMBL J01619, AAA23893.1;
CC EMBL M28989, AAA24616.1;
CC EMBL X00980, CAA25485.1;
CC EMBL AE000175, AAC73815.1;
CC EMBL D90711, BAA35388.1;
CC EMBL AE005250, AAG55045.1;
CC EMBL AB002552, BAB34169.1;
CC PIR: A28836; DEEC54.
CC
CC Ecogen: EG10933; sdhC.
CC InterPro: IPR00701; sdh_Cyt.
CC Pfam: PF01127; sdh_Cyt_1.
CC PROSITE: PS01000; SDH_CYT_1.
CC PROSITE: PS01001, SDH_CYT_2; 1.
CC Titratoxyllic acid cycle, Electron transport, Heme, 2200
CC Inner membrane; Complete proteome.
CC TRANSMEM 32 52
CC FT TRANSMEM 69 89 POTENTIAL.
CC FT TRANSMEM 109 129 POTENTIAL.
CC FT BINDING 40 30 HEME (POTENTIAL).
CC FT BINDING 84 84 HEME.
CC
CC SO SEQUENCE 129 AA; 1429 MW; 574950502545818 CRC64;

```

```

Query Match: 2.2% Score 7; DB 1; Length 12
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indel 0;
QY      33 LAGVLV 48
DB      123 LAGVLV 128

RESULT 4
DHSC_SALTY STANDARD: PRT: 129 AA.
ID DHSC_SALTY

```

08X140;
 15-JUN-2002 (Rel. 41, Created)
 15-JUN-2002 (Rel. 41, Last sequence update)
 15-JUN-2002 (Rel. 41, Last annotation update)
 Succinate dehydrogenase cytochrome b-556 subunit,
 SDHC OR STM0732 OR STY0775.
 08 Salmonella typhimurium, and
 09 Salmonella typhimurium.
 10 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 11 Salmonella.
 12 NCBI_TaxID=602, 601;
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563
 564
 565
 566
 567
 568
 569
 570
 571
 572
 573
 574
 575
 576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609
 610
 611
 612
 613
 614
 615
 616
 617
 618
 619
 620
 621
 622
 623
 624
 625
 626
 627
 628
 629
 630
 631
 632
 633
 634
 635
 636
 637
 638
 639
 640
 641
 642
 643
 644
 645
 646
 647
 648
 649
 650
 651
 652
 653
 654
 655
 656
 657
 658
 659
 660
 661
 662
 663
 664
 665
 666
 667
 668
 669
 670
 671
 672
 673
 674
 675
 676
 677
 678
 679
 680
 681
 682
 683
 684
 685
 686
 687
 688
 689
 690
 691
 692
 693
 694
 695
 696
 697
 698
 699
 700
 701
 702
 703
 704
 705
 706
 707
 708
 709
 710
 711
 712
 713
 714
 715
 716
 717
 718
 719
 720
 721
 722
 723
 724
 725
 726
 727
 728
 729
 730
 731
 732
 733
 734
 735
 736
 737
 738
 739
 740
 741
 742
 743
 744
 745
 746
 747
 748
 749
 750
 751
 752
 753
 754
 755
 756
 757
 758
 759
 760
 761
 762
 763
 764
 765
 766
 767
 768
 769
 770
 771
 772
 773
 774
 775
 776
 777
 778
 779
 780
 781
 782
 783
 784
 785
 786
 787
 788
 789
 790
 791
 792
 793
 794
 795
 796
 797
 798
 799
 800
 801
 802
 803
 804
 805
 806
 807
 808
 809
 810
 811
 812
 813
 814
 815
 816
 817
 818
 819
 820
 821
 822
 823
 824
 825
 826
 827
 828
 829
 830
 831
 832
 833
 834
 835
 836
 837
 838
 839
 840
 841
 842
 843
 844
 845
 846
 847
 848
 849
 850
 851
 852
 853
 854
 855
 856
 857
 858
 859
 860
 861
 862
 863
 864
 865
 866
 867
 868
 869
 870
 871
 872
 873
 874
 875
 876
 877
 878
 879
 880
 881
 882
 883
 884
 885
 886
 887
 888
 889
 890
 891
 892
 893
 894
 895
 896
 897
 898
 899
 900
 901
 902
 903
 904
 905
 906
 907
 908
 909
 910
 911
 912
 913
 914
 915
 916
 917
 918
 919
 920
 921
 922
 923
 924
 925
 926
 927
 928
 929
 930
 931
 932
 933
 934
 935
 936
 937
 938
 939
 940
 941
 942
 943
 944
 945
 946
 947
 948
 949
 950
 951
 952
 953
 954
 955
 956
 957
 958
 959
 960
 961
 962
 963
 964
 965
 966
 967
 968
 969
 970
 971
 972
 973
 974
 975
 976
 977
 978
 979
 980
 981
 982
 983
 984
 985
 986
 987
 988
 989
 990
 991
 992
 993
 994
 995
 996
 997
 998
 999
 1000

Matches 7: conservative 0: mismatches 0: indels 0: gaps 0:
 09 32 L1ACVAV 38
 111111
 Db 122 L1ACVAV 128
 RESULT 5
 OPA3_HUMAN
 ID OPA3_HUMAN STANDARD; PRT; 179 AA.
 AC C9H6K4;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Optic atrophy 3 protein.
 GN OPA3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Circulata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., DISFASIS, AND TISSUE SPECIFICITY.
 RX PubMed=11664429;
 FA Aiksterer V., Kleta P., Shaan A., Gahl W.A., Elpeleg O.,
 RT "Type III 3-methylglutaconic aciduria (optic atrophy plus syndrome, or
 RT Costoff optic atrophy syndrome): identification of the OPA3 gene and
 RT its founder mutation in Iraqi Jews."
 PI Am. J. Hum. Genet. 69:1218-1224(2001).
 EN [2]
 RP SEQUENCE FROM N.A.
 PA Matsubara K., Kamegai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
 PA Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,
 PA Nakamura Y., Iseida T., Sugano S.,
 RT "MEDO human cDNA sequencing project."
 RT Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 PC TISSUE-KIDNEY;
 RA Strausberg R.;
 RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: May play some role in mitochondrial processes.
 CC -1- TISSUE SPECIFICITY: ubiquitous. Most prominent expression in
 CC skeletal muscle and kidney.
 CC -1- DISEASE: Defects in OPA3 are the cause of type III 3-
 CC methylglutaconic aciduria (MGA type III); also known as optic
 CC atrophy plus syndrome or Costoff optic atrophy syndrome. MGA type
 CC III is a neuroophthalmologic syndrome consisting of early-onset
 CC bilateral optic atrophy and later-onset spasticity, extrapyramidal
 CC dysfunction, and cognitive deficit. Urinary excretion of 3-
 CC methylglutaconic acid and of 3-methylglutaric acid is increased.
 CC -1- SIMILARITY: BELONGS TO THE OPA3 FAMILY.
 CC
 CC This Swiss Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC use by non-profit institutions. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/annouce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DB EMBL: AK025843; EMBL525.1;
 DB EMBL: BC005059; AA005059.1;
 DB Genbank: NC0018142; OPA3.
 DB MIM: 606580;
 DB MIM: 258501;
 KW Mitochondrion; Coiled coil; Vision.
 FT DOMAIN 103 163 COILED COIL (POTENTIAL).
 SO SEQUENCE 179 AA: 1696 MW: 24948.948; 459AA; CP064;
 Query Match 2.2% Score 7: Db 1: Length 179;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 7: Conservative 0: Mismatches 0: Indels 0: Gaps 0;

QY 233 GENEAC 239
DB 157 GENEAC 163

RESULT 6

YPS1_YEAST STANDARD: PRT: 210 AA.
AC P36017;
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GTP-binding protein YPS1/YPS21.
GN YPS1 OR YPS21 OR YOK089C OR YOK3154C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetia; Saccharomycetaceae; Saccharomycetaceae.
OX NCBI_TaxID:4932;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE:94216368; PubMed:8163546,
RA Singer-Krueger H., Stenmark H., Duesterhoeft A., Philippson P.,
RA Yoo J.-S., Gallwitz D., Zerial M.;
RT *Role of three rab5-like GTPases, Yps1p, Yps2p, and Yps3p, in the
RT endocytic and vacuolar protein sorting pathways of yeast.*;
RL J. Cell Biol. 125:283-298(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE:94185635; PubMed:8137814;
RA Horadzowsky B.F., Busch G.R., Emr S.D.;
RT *YPS21 encodes a rab5-like GTP binding protein that is required for
RT the sorting of yeast vacuolar proteins.*;
RL EMBO J. 13:1297-1309(1994).
RN [3]
RP SEQUENCE FROM N.A.
RA MEDLINE:97344368; PubMed:9200815;
RA Voss H., Benes V., Andrade M.A., Valencia A., Teodoro C.,
RA Schwager C., Paces V., Sander C., Ansorge W.;
RT *DNA sequencing and analysis of 130 kb from yeast chromosome XV.*;
RL Yeast 13:655-672(1997).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF 1-174.
RA MEDLINE:20221439; PubMed:1376108,
RA Esters H., Alexandrov K., Constantinescu A.-T., Grody P.S.,
RA Schofield A.J.;
RT *High-resolution crystal structure of S. cerevisiae Yps1(DeltaC15)-
RT GppNHp, a small GTP-binding protein involved in regulation of
RT endocytosis.*;
RL J. Mol. Biol. 298:111-121(2000).
RN [5]
RP FUNCTION: REQUIRED FOR TRANSPORT IN THE ENDOCYTIC PATHWAY AND
RP FOR CORRECT SORTING OF THE VACUOLAR HYDROLASES SUGGESTING A
RP POSSIBLE INTERSECTION OF THE ENDOCYTIC WITH THE VACUOLAR SORTING
RP PATHWAY. ACT EITHER AT THE STAGE OF DELIVERY OF PROTEINS FROM THE
RP GOLGI TO THE LATE ENDOSOME OR FROM THE LATE ENDOSOME TO THE
RP VACUOLE.
CC -! SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC RAB5 HOMOLOG.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL contribution.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by any other
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X76171; CAA53769.1;
CC EMBL: Z29338; CAA82543.1;
CC EMBL: X94335; CAA64010.1;
CC EMBL: Z74997; CAA99285.1;
CC PIR: S40001; S40001.
CC PIR: S43399; S43399.
CC PDB: 1EK0; 17-APR-00.

DR Sbj: S0056515; YPS21.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_trimming.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; Ras_1.
DR PRINTS: PR00449; PASTIN/FRMN.
DR SMART: SM00175; Rab_1.
DR TIGRFAMs: TIGR00231; small_gtp.
KW Protein transport; GTP-binding; Lipoprotein; Prenylation;
KW 3D-structure.
FT NP_BIND 14 21 GTP.
FT NP_BIND 62 66 GTP.
FT NP_BIND 120 123 GTP.
FT DOMAIN 36 44 EFECTOR REGION (POTENTIAL)
FT LIPID 208 208 GERANYL-GERANYL (BY SIMILAR)
FT LIPID 210 210 GERANYL-GERANYL (BY SIMILAR)
FT MUTAGEN 21 21 DEFECT.
FT
SO SEQUENCE 210 AA; 23081 MW; 4477AD090C3F09 C6764;

Query Match 2.2% Score 7; DB 1; Length 233
Best Local Similarity 100.0% Prod. No. 209
Matches 7; Conservative 6; Mismatches 0; Indels

QY 233 KTAEEON 239
DB 174 KTAEEON 180

RESULT 7

Y282_TREPA STANDARD: PRT: 219 AA.
AC 083306;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TP0282.
GN TP0282.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID:160;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=Nichols;
RX MEDLINE:96332770; PubMed:9445876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sult
RA Fraser R., Galin M., Hickey R., Gargano P., Kacham S.,
RA Sodergren E., Hardam J.M., McLeod M.P., Salzberg S., Bo
RA Khakh H., Richardson D., Howell J.K., Chidambaram M., H
RA McDonald L., Artach P., Bowman P., Nelson M.D., Fujita
RA Hatch R., Horst K., Roberts K., Sridhary M., Weidman C.
RA Venter J.C.;
RT *Complete genome sequence of Treponema pallidum, the sy
RT Spirochete.*;
RL Science 281:1375-1388(1998).
CC -! SIMILARITY: CONTAINS 2 TPR REPEATS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through
CC between the Swiss Institute of Bioinformatics and the
CC the European Bioinformatics Institute. There are no res
CC use by non-profit institutions as long as its content
CC modified and this statement is not removed. Usage by
CC entities requires a license agreement (See [http://www.isb.ch/](http://www.isb.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE001209; AAC65279.1;
CC DR TIGR: TP0282.
DR InterPro: IPR001440; TPR.
KW Hypothetical protein; Repeat; TPR repeat; Transmembrane
KW Complete proteome.
FT TRANSMEM 21 43 P-IDENTICAL
FT REPEAT 112 145 TPR 1
FT REPEAT 149 182 TPR 2

Matches 7: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

OY 246 EONFLQ 242
DB 215 EONFLQ 221

RESULT 10

PROB_PASMO
ID PROB_PASMO STANDARD: PRT: 367 AA.
AC U9CJ05:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glutamate 5-kinase (EC 2.7.2.11) (gamma-glutamyl kinase) (CK).
GN PROB OR PM1896.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_Taxid=747;
RN [1]

SEQUENCE FROM N.A.

RC STRAIN-Pm70:
RA MEDLINE=2145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whitam T.S., Kapur V.,
RT *Complete genomic sequence of Pasteurella multocida pm70.*;
RL Proc Natl Acad Sci U S A 98:3460-3465(2001).

CC -1- FUNCTION: Catalyzes the transfer of a phosphate group to glutamate
to form glutamate 5-phosphate which rapidly cyclizes to 5-
oxoproline.
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate = ADP + L-glutamate 5-
phosphate.

CC -1- PATHWAY: Proline biosynthesis: first step.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -1- SIMILARITY: BELONGS TO THE GLUTAMATE 5-KINASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL: AE006227; AAK03980.1; ALT_INIT.

DR InterPro: IPR001048; Aa_kinase.

DR InterPro: IPR002478; PUA.

DR Pfam: PF00696; aakkinase; 1.

DR Pfam: PF01472; PUA; 1.

DR PRINTS: PR00474; GLO5KINASE.

DR SMART: SM00359; PUA; 1.

DR TIGRFAMs: TIGR01027; PROB; 1.

DR PROSITE: PS00903; GLUTAMATE_5_KINASE; 1.

DR TRANSFERASE_KINASE; Proline biosynthesis: Complete proteome.

SU SEQUENCE 367 AA: 39894 MW: E62DA632E78FE5 CRC64;

Query Match 2.2%; Score 7; DB 1; Length 367.

Best Local Similarity 100.0%; Pred. No. 33;
Matches 7: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

OY 47 LVALLVQ 43
DB 155 LVALLVQ 161

RESULT 11

PROB_HAEIN

ID PROB_HAEIN STANDARD: PRT: 368 AA.

AC P43763:
DT 01-NOV-1995 (Rel. 42, Created)

DT 01-NOV-1995 (Rel. 42, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Glutamate 5-kinase (EC 2.7.2.11) (gamma-glutamyl kinase)

GN PROB OR H10900.

OS Haemophilus influenzae.

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurella

OC Haemophilus

OX NCBI_Taxid=727;
RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-Pd / KW20 / ATCC 51907;

RA MEDLINE=95350630; PubMed=7542800;

RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., K

RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., M

RA McKenney K., Sifton G., Fitzhugh W., Fries C.A., Gocayne

RA Scott J.D., Shirley B., Liu T.-I., Glodek A., Kelley J.M

RA Widmann J.F., Phillips C.A., Spriggs T., Heidelberg E., C

RA Ufferkamp T.R., Hanna M.C., Nguyen B.T., Saudek D.M., B

RA Fine L.D., Fritchman J.L., Fritchman J.L., Goodgen N.S.

RA Guehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smi

RA Venter J.C.;
RT "Whole genome random sequencing and assembly of Haemophi

RL Science 269:496-512(1995).

CC -1- FUNCTION: Catalyzes the transfer of a phosphate group
to form glutamate 5-phosphate which rapidly cyclizes
to 5-oxoproline.

CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate = ADP + L-glut

CC -1- PATHWAY: Proline biosynthesis: first step.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -1- SIMILARITY: BELONGS TO THE GLUTAMATE 5-KINASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced throug

CC between the Swiss Institute of Bioinformatics and the

CC the European Bioinformatics Institute. There are no r

CC use by non-profit institutions as long as its content

CC modified and this statement is not removed. Usage by

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL: U03772; M022560.1; -

DR TIGR: H10900; -

DR InterPro: IPR001048; Aa_kinase.

DR InterPro: IPR001057; Glu_5kinase.

DR InterPro: IPR002478; PUA.

DR Pfam: PF00696; aakkinase; 1.

DR Pfam: PF01472; PUA; 1.

DR PRINTS: PR00474; GLO5KINASE.

DR SMART: SM00359; PUA; 1.

DR TIGRFAMs: TIGR01027; PROB; 1.

DR PROSITE: PS00903; GLUTAMATE_5_KINASE; 1.

DR TRANSFERASE_KINASE; Proline biosynthesis: Complete prote

SU SEQUENCE 368 AA: 40134 MW: E32684A23709631 CRC64;

Query Match 2.2%; Score 7; DB 1; Length 368

Best Local Similarity 100.0%; Pred. No. 33;
Matches 7: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

OY 37 LVALLVQ 43
DB 154 LVALLVQ 160

RESULT 12

DXR_FUSNN

ID DXR_FUSNN STANDARD: PRT: 390 AA.

AC Q8R622:
DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE 1-deoxy-D-xylulose 5-phosphate reductoisomerase (EC 1.1.1.1)
DE reductoisomerase (1-deoxyxylulose 5-phosphate reductio
GN DXR OR FN1324.
OS Fusobacterium nucleatum (subsp. nov.) (outum).

```

CC Bacteria; Fusobacteriales; Fusobacterium
CC NCBI_TaxID=76866;
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN ATCC 25586;
CC MEDLINE=2186444; PubMed=11864104;
CC Kapatral V., Anderson I., Tadmor N., Frenkel A.,
CC Bhattacharya A., Bartman A., Gardner W., Greekin G., Zhu L.,
CC Vastava O., Chu L., Kogan Y., Chada O., Goldsmn E., Hernal A.,
CC Larsen N., D'Souza M., Malinas T., Pusch G., Haselkorn R.,
CC Fournier M., Kyprides N., Overbeck R.;
CC "Genomic sequence and analysis of the oral bacterium Fusobacterium
CC nucleatum strain ATCC 25586."
CC J. Bacteriol. 184:2005-2014(2002)
CC -1- FUNCTION: Catalyzes the NADP-dependent rearrangement and reduction
CC of 1-deoxy-D-xylulose-5-phosphate (DXR) to 2-C-methyl D-erythritol
CC 4-phosphate (MEP) (by similarity).
CC -1- CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate + NADP(+)
CC = 1-deoxy-D-xylulose 5-phosphate + NADPH.
CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; second
CC step.
CC -1- SIMILARITY: BELONGS TO THE DXR FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: AE010637; AAL95520.1;
CC 1SEQUENCE BIOSYNTHESIS; oxidoreductase; NADP; Complete proteome.
CC NP_BIND 10 NADPH (POTENTIAL).
CC SEQUENCE 390 AA; 43964 MW; 1C3771487885D3ED C6C64;
CC
CC Query Match: 2.28, Score 7, DB 1, Length 390;
CC Best Local Similarity 100.0%; Pred. No. 45;
CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 56 SP0DAIV 62
CC |||||
CC 61 SP0DAIV 67
CC
CC RESULT 13
CC YCWM_ECOLI
CC ID YCWM_ECOLI STANDARD; PRI: 396 AA.
CC AC P75876; Q98702;
CC DT 15-JUN-1998 (Rel. 35, created)
CC DT 16-OCT-2001 (Rel. 40, last sequence update)
CC DT 16-OCT-2001 (Rel. 40, last annotation update)
CC DE Hypothetical protein ycw.
CC GN YCWM_OP P09667
CC OS Escherichia coli.
CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC Escherichia
CC CC [1]
CC NCBI_TaxID=562;
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN-K12 / MG1655;
CC MEDLINE=97426617; PubMed=9778503;
CC Rhatner F.R., Plunkett G., III, Bloch C.A., Perna N.T., Putland V.,
CC Riley M., Colado Vides J., Glasner J.B., Rode C.K., Mayhew G.P.,
CC Gregor J., Davis N.W., Kirkpatrick B.A., Goeden M.A., Rose T.T.,
CC Mau B., Shao Y.;
CC "The complete genome sequence of Escherichia coli K-12."
CC Science 277:1453-1474(1997).
CC [2]
CC SPOURCE FROM N.A.
CC STRAIN-K12;
CC MEDLINE 97061202; PubMed=8905232;
CC Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,

```

```

CC 1Kmeto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
CC Kimura G., Kitagawa M., Matsuo K., Nasais S., Miki T., Mizobuchi K.,
CC Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
CC Sampei G., Seki Y., Takami H., Takemoto K., Wada G., Yamamoto Y.,
CC Yano M., Horinouchi T.;
CC "A 718-kt DNA sequence of the Escherichia coli K-12 genome
CC corresponding to the 3,272,946 nt region on the linkage map."
CC J. Bacteriol. 183:137-155(1996).
CC [1]
CC -1- SIMILARITY: BELONGS TO THE DEPO6064 FAMILY
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: A000196; AAC74053.1; ALT_INIT.
CC EMBL: 090744; AAC45732.1;
CC ProGene: E013725; YCWM.
CC DR InterPro: IPR002478; YCWM.
CC DR InterPro: IPR001250; Prenyl_site.
CC DR InterPro: IPR000051; SAM_bind.
CC DR SMART: SM00359; PUA; 1.
CC KW Hypothetical protein; Complete proteome.
CC SEQUENCE 496 AA; 44446 MW; 0905FCEB62734094 C6C64;
CC
CC Query Match: 2.28, Score 7, DB 1, Length 396;
CC Best Local Similarity 100.0%; Pred. No. 36;
CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 23 LVGLDLS 29
CC |||||
CC 124 LVGLDLS 130
CC
CC RESULT 14
CC PTMB_BACST
CC ID PTMB_BACST STANDARD; PRI: 471 AA.
CC AC P50852;
CC DT 01-OCT-1996 (Rel. 34, created)
CC DT 01-OCT-1996 (Rel. 34, last sequence update)
CC DT 15-JUN-2002 (Rel. 41, last annotation update)
CC DE Pts system, mannitol-specific TIBC component (EIIIC-Mtl) (Mannitol-
CC dephosphorylase TIBC component) (Phosphotransferase enzyme II, BC component)
CC DE (EC 2.7.1.69) (EII-Mtl).
CC GN MTLA.
CC OS Bacillus stearothermophilus.
CC Bacteria; Firmicutes; Bacillales; Geobacillus.
CC CC [1]
CC NCBI_TaxID=1422;
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN ATCC 7954;
CC MEDLINE=96421984; PubMed=8824601;
CC Benstrin S.A., Tolner R., Howe Furkens P.H., Konings W.N.,
CC Reddihall G.T.;
CC "Cloning, expression, and isolation of the mannitol transport protein
CC from the thermophilic bacterium Bacillus stearothermophilus."
CC J. Bacteriol. 178:5586-5591(1996).
CC -1- FUNCTION: THIS IS A COMPONENT OF THE PHOSPHOTRANSFER-DEPENDENT
CC SUGAR PHOSPHOTRANSFER SYSTEM (PTS). A MAJOR CARBOHYDRATE ACTIVE
CC -TRANSPORT SYSTEM. THE TIBC DOMAIN CONTAINS THE SUGAR BINDING SITE
CC AND THE TRANSMEMBRANE CHANNEL. THE IIA DOMAIN CONTAINS THE PRIMARY
CC PHOSPHORYL GROUP TO THE IIB DOMAIN WHICH FINALLY TRANSFERS IT TO
CC THE SUGAR.
CC -1- CATALYTIC ACTIVITY: protein N-phosphotransferase + sugar -> protein
CC N-phosphotransferase + sugar phosphate.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: CONTAINS 1 PTS IIB DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 PTS EIIIC DOMAIN.
CC

```

CC This SWISS-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed, resays by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U18943; AAC44463.1; .
 DR InterPro: IPR004718; PISLIC_m1A
 DR TricFAMS: TRICR00851; m1A: 1.
 KW Phosphotransferase system: Sugar transport: Transferase:
 KW Transmembrane; Phosphorylation.
 FT DOMAIN 1 339
 FT DOMAIN 1 339
 FT TRANSMEM 30 50
 FT TRANSMEM 54 74
 FT TRANSMEM 92 112
 FT TRANSMEM 143 163
 FT TRANSMEM 219 239
 FT TRANSMEM 273 293
 FT TRANSMEM 294 314
 FT TRANSMEM 319 339
 FT MOD_RES 262 282
 FT MOD_RES 389 389
 SO SEQUENCE 471 AA; 50153 MW; 3BF2F93C8ED968C CRC64;

Query Match 2.28; Score 7; DB 1; Length 471;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 33 LAGVLA 39
 DDb 422 LAGVLA 328

RESULT 15
 DHAL_ASPNG STANDARD; PRT: 497 AA.
 AC P41751;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Aldehyde dehydrogenase (Ec 1.2.1.3) (ALDH).
 GN ALDA.
 OS Aspergillus niger.
 CC Eukaryota, Fungi, Ascomycota, Pezizomycotina, Eurotiomycetes;
 CC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 CC NCBI_TaxID=5061;
 CC [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=90108706; PubMed=2606157;
 RA O'Connell M.J., Kelly J.M.;
 RT *Physical characterization of the aldehyde-dehydrogenase-encoding
 RT gene of Aspergillus niger.*;
 RL Gene 84-173-180(1983)
 CC -1- CATALYTIC ACTIVITY: An aldehyde + NAD(+) + H(2)O = an acid + NADH.
 CC -1- PATHWAY: Ethanol utilization; second step.
 CC -1- SIMILARITY: BELONGS TO THE ALDEHYDE DEHYDROGENASES FAMILY.
 CC -----
 CC This SWISS-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed, usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M32531; AAA87596.1; .
 DR HSSP: P05091; 1CW3.
 DR InterPro: IPR002086; Aldehyde_dehydr.
 DR Pfam: PF00171; aldehyd. 1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS: 1.

DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU: 1.
 KW Oxidoreductase; NAD.
 FT NP_BIND 242 247
 FT ACT_SITE 265 265
 FT ACT_SITE 294 294
 FT ACT_SITE 394 394
 SO SEQUENCE 497 AA; 53809 MW; 4DCA454FDDA50E8 CRC64;

Query Match 2.28; Score 7; DB 1; Length 497;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 7; Conservative 0; Mismatches 0; Indels 0;

UY 230 VIKTAE 236
 DDb 186 VIKTAE 192

RESULT 16
 CSU2_DROME STANDARD; PRT: 501 AA.
 AC U9VM16;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable cytochrome P450 2b4 (Ec 1.14.14.1) (CYP2B4).
 GN CYP2B4 OR C6081.
 OS Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea;
 CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; 16
 CC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 CC NCBI_TaxID 7227;
 CC [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 KX MEDLINE=20194006; PubMed=10741192;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gonsky
 RA Apanatides P.C., Scherer S.E., Li P.W., Hoskins R.A., G
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Heu
 RA Sutton G.G., Wortman J.P., Yandell M.D., Zhang Q., Chen
 RA Randon R.C., Rogers Y.-H.C., Rhee B.G., Champe M., Pe
 RA Wan K.H., Doyle C., Baxter F.C., Holt R.A., Nelson G.P., B
 RA Abell J.F., Abmayyan A., An H.-J., Andrews-Pfannkoch C.
 RA Bailey R.M., Basu A., Baxendale J., Bayraktaroglu L., B
 RA Beeson K.Y., Beus P.V., Bernier R.P., Bhandari D., Bol
 RA Botkova P., Botchan M.R., Buck J., Buckstein P., Bro
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davi
 RA de Pablo R., Delcher A., Dew Z., Mays A.D., Dew I., D
 RA Dodson K., Doup L.E., Downes M., Dudan-Kocha S., Dunko
 RA Durbin K.J., Evangelista C.C., Forier C., Ferreira S.,
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., G
 RA Gluck A., Gong F., Gottlieb J.H., Gu Z., Guan P., Hart
 RA Harris N.L., Harvey D., Heiman J.J., Hernandez J.R., H
 RA Hestlin D., Houston K.A., Howland J.W., Wei M.-H., Ilek
 RA Jaitai M., Kalush F., Karpas A., Ke Z., Kennedy J.A.
 RA Kimmel B.F., Kodira G.D., Kraft C., Kravitz S., Kulp D.
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y.
 RA Liu X., Mattei B., McIntosh T.C., Melrod M.P., Mephor
 RA Mekurov G., Milshina N.V., Mobury C., Morris J., Mos
 RA Mount S.M., Moy M., Murphy B., Murphy L., Murty D.M., M
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pace
 RA Palczuk D.M., Pittman G.S., Fan S., Pollard J.D., Puri V
 RA Reiter K., Remington K., Saunders R.D., Scheeler P.
 RA Shie R.C., Siden-Kiamos I., Simpson M., Skupski M., F
 RA Spletter E., Spradling A.C., Stapleton M., Strong R., Su
 RA Svidersky R., Tector C., Turner R., Venter E., Wang A.
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach
 RA Williams S.M., Woodage T., Worley K.F., Wu D., Yang S
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT *The genome sequence of Drosophila melanogaster.*;
 RL Science 287:2185-2195(2000).
 RN [1]

```

4P CONCEPTUAL TRANSLATION.
4A Nelson B.:
4L Unpublished observations (SEP-2000).
4C -1- FUNCTION: MAY BE INVOLVED IN THE METABOLISM OF INSECT HORMONES AND
4C IN THE BREAKDOWN OF SYNTHETIC INSECTICIDES (BY SIMILARITY).
4C -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) -> RH +
4C oxidized flavoprotein + H(2)O.
4C -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum
4C (potential).
4C -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY
4C -1- CAUTION: Ref.1 sequence differs from that shown due to erroneous
4C gene model prediction.
4C This SWISS-PROT entry is copyright. It is produced through a collaboration
4C between the Swiss Institute of Bioinformatics and the EMBL outstation -
4C the European Bioinformatics Institute. There are no restrictions on its
4C use by non-profit institutions as long as its content is in no way
4C modified and this statement is not removed. Usage by and for commercial
4C entities requires a license agreement (See http://www.isb-sib.ch/announce/
4C or send an email to license@isb-sib.ch).
4C -----
4C EMBL: AE003609; AF52235.1; ALT_SEQ.
4C FlyBase: FBgn001668; Cyp2b2
4C InterPro: IPR01128; Cytochrome_P450.
4C Pfam: PF00067; P450_1.
4C PRINTS: PR00385; P450.
4C PROSITE: PS00086; CYTOCHROME_P450_1.
4C Oxidoreductase; Monooxygenase; Membrane; Heme; Microsome;
4C Endoplasmic reticulum; Hypothetical protein.
4C BINDING: 446 446 HEME (BY SIMILARITY).
4C SEQUENCE: 501 AA; 5822 MW; 9A9BFF34930A5E CIRCULAR.
4C -----
4C Query Match: 2.28; Score 7; DB 1; Length 501;
4C Best Local Similarity: 100.0%; Pred. No. 44;
4C Matches: 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
4C -----
4C 88 LVLVAAA 94
4C |||||
4C 453 LVLVAAA 459
4C -----
4C RESULT 17
4C TM02_POPK1
4C AC Q43054; 064949; Q40906; Q4044;
4C 15-DEC-1998 (Ref. 37, Created)
4C 15-DEC-1998 (Ref. 37, Last sequence update)
4C 15-JUN-2002 (Ref. 41, Last annotation update)
4C Trans-cinnamate 4-monooxygenase (EC 1.14.13.11) (Cinnamic acid
4C 4-hydroxylase) (C4H) (P450C4H) (Cytochrome P450 73).
4C CYP73A16 OR CYP73.
4C OS Populus kirkamiansis (Aspen).
4C Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
4C Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
4C Eurosids I; Malpighiales; Salicaceae; Populus.
4C NCBI_TaxID=3698;
4C [1]
4C SEQUENCE FROM N.A.
4C RX MEDLINE:97141312; PubMed:8987656;
4C Kawai S., Mori A., Shiohara T., Kajita S., Katayama Y., Morishashi N.:
4C "Isolation and analysis of cinnamic acid 4-hydroxylase homologous
4C genes from a hybrid aspen, Populus kirkamiansis";
4C Biosci. Biotechnol. Biochem. 60:1586-1597(1996).
4C -1- FUNCTION: CONTROLS CARBON FLUX TO PIGMENTS ESSENTIAL FOR
4C POLLINATION OR UV PROTECTION, TO NUMEROUS PHYTOALEXINS (2H)
4C BY PLANTS WHEN CHALLENGED BY PATHOGENS, AND TO LIGNINS.
4C -1- CATALYTIC ACTIVITY: Trans-cinnamate + NADPH + O(2) -> 4-
4C hydroxycinnamate + NADP(+) + H(2)O.
4C -1- PATHWAY: Phenylpropanoid metabolism; second step.
4C -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
4C This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

4C between the Swiss Institute of Bioinformatics and the EMBL outstation -
4C the European Bioinformatics Institute. There are no restrictions on its
4C use by non-profit institutions as long as its content is in no way
4C modified and this statement is not removed. Usage by and for commercial
4C entities requires a license agreement (See http://www.isb-sib.ch/announce/
4C or send an email to license@isb-sib.ch).
4C -----
4C EMBL: D82812; BA011576.1;
4C DB EMBL: D82815; BA011579.1;
4C InterPro: IPR01128; Cytochrome_P450.
4C Pfam: PF00067; P450_1.
4C PRINTS: PR00385; P450.
4C PROSITE: PS00086; CYTOCHROME_P450_1.
4C Oxidoreductase; Monooxygenase; Heme; NADP.
4C BINDING: 447 447 HEME (BY SIMILARITY).
4C SEQUENCE: 505 AA; 57973 MW; FA970FA9B138151 CIRCULAR.
4C -----
4C Query Match: 2.28; Score 7; DB 1; Length 505;
4C Best Local Similarity: 100.0%; Pred. No. 44;
4C Matches: 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
4C -----
4C 36 LVLVALLY 42
4C |||||
4C 17 LVLVALLY 23
4C -----
4C RESULT 18
4C TMA_LYCES
4C ID TMA_LYCES
4C AC Q22478;
4C 15-DEC-1998 (Ref. 36, Created)
4C 15-DEC-1998 (Ref. 37, Last sequence update)
4C 16-MAY-2001 (Ref. 46, Last annotation update)
4C Importin alpha subunit (tau)protein alpha subunit (KAR alpha).
4C Lycopersicon esculentum (Tomato).
4C Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
4C Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
4C Asteridae; easterids I; Solanales; Solanaceae; Solanum.
4C NCBI_TaxID=4081;
4C [1]
4C SEQUENCE FROM N.A.
4C Kuehl T., Murrer L., Citovsky V., Catal Y.:
4C Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases
4C -1- FUNCTION: BINDS SPECIFICALLY AND DIRECTLY TO SUBSTRATES CONTAINING
4C EITHER A SIMPLE OR BIPHARTITE NLS MOTIF. PROMOTES DOCKING OF IMPORT
4C SUBSTRATES TO THE NUCLEAR ENVELOPE. SEEMS TO ACT AS A CYTOSOLIC
4C PROTECTOR FOR BOTH SIMPLE AND BIPHARTITE NLS MOTIFS (BY SIMILARITY).
4C -1- SUBUNIT: FORMS A COMPLEX WITH IMPORTIN BETA-1 SUBUNIT (BY
4C SIMILARITY).
4C -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
4C -1- SIMILARITY: BELONGS TO THE IMPORTIN ALPHA FAMILY.
4C -1- SIMILARITY: CONTAINS 8 ARM REPEATS.
4C This SWISS-PROT entry is copyright. It is produced through a collaboration
4C between the Swiss Institute of Bioinformatics and the EMBL outstation -
4C the European Bioinformatics Institute. There are no restrictions on its
4C use by non-profit institutions as long as its content is in no way
4C modified and this statement is not removed. Usage by and for commercial
4C entities requires a license agreement (See http://www.isb-sib.ch/announce/
4C or send an email to license@isb-sib.ch).
4C -----
4C EMBL: AF017552; AC23722.1;
4C DB HSSP: G02821; 1B85.
4C InterPro: IPR002525; ARMAD_110.
4C InterPro: IPR002522; Importina_B.
4C Pfam: PF00514; Armadillo_seq_8.
4C PRINTS: PR01749; IBB_1.
4C SMART: SM00185; ARM_8.
4C PROSITE: PS0176; ARM_REPEAT; 5.
4C TransPort: PTC0176; Importin transport; Repeat.
4C P-MAIN 12 51
4C REPEAT 106 151 ARM 1.
4C REPEAT 152 196 ARM 2.

```

FT REPEAT 197 234 ARM 3.
 FT REPEAT 235 279 ARM 4.
 FT REPEAT 280 319 ARM 5.
 FT REPEAT 320 362 ARM 6.
 FT REPEAT 363 403 ARM 7.
 FT REPEAT 404 445 ARM 8.
 FT REPEAT 446 527 ARM 9.
 FT DOMAIN 446 527 ASP/GDU-RICH (ACIDIC).
 NO SEQUENCE 527 AA: 58605 MW: 443601641CFF4817 CRR64;
 Query Match 2.28; Score 7; DB 1; Length 527;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 18 LGICALV 24
 1111111
 198 LGICALV 204
 RESULT 19
 LAC1_PHLA STANDARD; PRT: 548 AA.
 ID LAC1_PHLA
 AC Q01679;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Laccase precursor (PF 1 to 3 2) (Benzoinoligoxigen oxidoreductase)
 DE (Urishiol oxidase) (Ligninolytic phenoloxidase)
 GN LAC.
 OS Phlebia radiata (White-rot fungus).
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
 OC Apophloeales; Corticiaceae; Phlebia.
 CX NCBI_TaxID=5308;
 RN [1]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE
 RC STRAIN=ATCC 64658;
 RX MEDLINE=92065223; PubMed 1955850.
 RA Saloheimo M., Niku-Paavola M.L., Knowles T.K.;
 RT "Isolation and structural analysis of the laccase gene from the
 RT lignin-degrading fungus *Phlebia radiata*."
 RI J. Gen. Microbiol. 137:1547-1544(1991).
 CC
 CC -1- FUNCTION: MOST PROBABLY PLAYS AN IMPORTANT ROLE IN LIGNIN
 CC DEGRADATION
 CC -1- CATALYTIC ACTIVITY: 4 benzenediol + O(2) = 4 benzenesemiquinone + 2
 CC H(2)O.
 CC -1- COFACTOR: BINDS 4 CU-IONS PER MOLECULE. THREE DISTINCT CU
 CC CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE
 CC 3 OR COUPLED BINUCLEAR.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES
 CC -1- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X52134; CAA36379.1;
 CC EMBL: A20705; CAA01513.1;
 CC PIR: S18746; S18746.
 CC HSSP: P37064; IASO.
 CC InterPro: IPR001117; Cu-oxidase.
 CC InterPro: IPR002355; Multicu_oxidase2.
 CC Pfam: PF00394; Cu-oxidase2.
 CC PROSITE: PS00079; MULTICOPPER_OXIDASE1; 1.
 CC PROSITE: PS00080; MULTICOPPER_OXIDASE2; 1.
 CC Oxidoreductase; Signal; Copper; Metal-binding; Lignin degradation.
 KW Glycoprotein; Repeat.
 FT SIGNAL 1 21
 FT SIGNAL 22 548 LACCASE.
 FT SIGNAL 22 548 PLASTOCYANIN-LIKE 1.
 FT DOMAIN 22 148

FT DOMAIN 160 304 PLASTOCYANIN-LIKE 2.
 FT DOMAIN 373 496 PLASTOCYANIN-LIKE 3.
 FT METAL 87 85 COPPER (TYPE 2) (PROBABLE).
 FT METAL 87 87 COPPER (TYPE 3) (PROBABLE).
 FT METAL 130 130 COPPER (TYPE 4) (PROBABLE).
 FT METAL 132 132 COPPER (TYPE 3) (PROBABLE).
 FT METAL 132 132 COPPER (TYPE 3) (PROBABLE).
 FT METAL 418 418 COPPER (TYPE 2) (PROBABLE).
 FT METAL 421 421 COPPER (TYPE 2) (PROBABLE).
 FT METAL 423 423 COPPER (TYPE 3) (PROBABLE).
 FT METAL 423 423 COPPER (TYPE 3) (PROBABLE).
 FT METAL 474 474 COPPER (TYPE 1) (PROBABLE).
 FT METAL 474 474 COPPER (TYPE 1) (PROBABLE).
 FT METAL 475 475 COPPER (TYPE 3) (PROBABLE).
 FT METAL 479 479 COPPER (TYPE 1) (PROBABLE).
 FT METAL 484 484 COPPER (TYPE 1) (PROBABLE).
 FT CARBOHYD 75 75 N-LINKED (GLCNAC...) (POST-5
 FT CARBOHYD 352 352 N-LINKED (GLCNAC...) (POST-5
 FT CARBOHYD 402 402 N-LINKED (GLCNAC...) (POST-5
 FT CARBOHYD 457 457 N-LINKED (GLCNAC...) (POST-5
 NO SEQUENCE 548 AA: 58601 MW: 44567723056B2 CRR64;
 Query Match 2.24; Score 7; DB 1; Length 54;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 7; Conservative 0; Mismatches 0; Indels 0;
 DB 308 DVDNTW1 314
 1111111
 275 DVDNTW1 281
 RESULT 20
 PRIZ_DROVI STANDARD; PRT: 584 AA.
 ID PRIZ_DROVI
 AC Q24760;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Friczled protein precursor (Friczled-1) (Dz1).
 GN FZ.
 OS Drosophila virilis (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; M
 OC Muscomorpha; Ephydroidea; Erosophiliidae; Drosophila.
 CX NCBI_TaxID=7244;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=ovary;
 RX MEDLINE=96366421; PubMed=8770598;
 RA Jones K.H., Liu J., Adler P.N.;
 RT "Molecular analysis of EMS induced friczed mutations in
 RT melanogaster".
 RI Genetics 142:205-215(1996).
 CC
 CC -1- FUNCTION: Receptor for Wnt proteins. Most of friczed
 CC are coupled to the beta-catenin canonical signaling
 CC leads to the activation of dishevelled proteins. It
 CC GSK-3 kinase, nuclear accumulation of beta-catenin
 CC of Wnt target genes. A second signaling pathway inv
 CC calcium fluxes has been seen for some family member.
 CC not yet clear if it represents a distinct pathway
 CC integrated in the canonical pathway, as PKC seems t
 CC for Wnt mediated activation of GSK-3 kinase. Both
 CC to involve interactions with G-proteins. Regulated
 CC the cytoskeletons of epidermal cells to produce a l
 CC of cuticular hairs and bristles.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (P
 CC -1- DOMAIN: Lys-Thr-X-X-X-Tip motif is involved in the
 CC the Wnt/beta-catenin signaling pathway (by similar
 CC similarity). The fz domain is involved in binding with a
 CC -1- SIMILARITY: BELONGS TO FAMILY FZ/SW/OF G PROTEIN
 CC RECEPTORS.
 CC -1- SIMILARITY: CONTAINS 1 FRICZLED (FZ) DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced th

GN
MICA.
OS
Bacillus subtilis.
OT
Bacteria: Firmicutes, Bacillales, Bacillaceae; bacillus.
OX
NCBI_TaxID=1423;
RN
[1]

```
CC or send an email to license@lab-sib.ch.
CC
EMBL: 036161; BAA07350.1; .
DR
EMBL: D50453; BAA09029.1; .
DR
EMBL: Z9106; CAU12206.1; .
```


DR HSSP: P00550; 1A3A.
 DR Subtilisin; BGI1215; mLA.
 DR InterPro: IPR004718; PSLIC_mLA.
 DR Pfam: PF00359; PTS_E1A_2; 1.
 DR ProDom: PD001689; PTS_E1A_2; 1.
 DR TricFAMS: TricF00851; mLA; 1.
 DR PROSITE: PS00372; PTS_E1A_2; 1.
 KW Phosphotransferase system: Sugar transport: Transferase:
 Transmembrane; Phosphorylation; Complete proteome.
 FT DOMAIN 1 339 ELIC DOMAIN.
 FT DOMAIN 340 495 ELIB DOMAIN.
 FT DOMAIN 495 610 ELIA DOMAIN.
 FT TRANSMEM 30 50 POTENTIAL.
 FT TRANSMEM 54 74 POTENTIAL.
 FT TRANSMEM 90 110 POTENTIAL.
 FT TRANSMEM 141 161 POTENTIAL.
 FT TRANSMEM 167 187 POTENTIAL.
 FT TRANSMEM 189 209 POTENTIAL.
 FT TRANSMEM 219 239 POTENTIAL.
 FT TRANSMEM 251 271 POTENTIAL.
 FT TRANSMEM 290 310 POTENTIAL.
 FT TRANSMEM 319 339 POTENTIAL.
 FT MOD_RES 262 262 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 547 547 PHOSPHORYLATION (BY SIMILARITY).
 FT CONFLICT 134 135 ML -> IR (IN REF.1).
 SV SEQUENCE 610 AA; 65426 MW; E446F36D50E5A907 CRC64;

Query Match 2.28; Score 7; DB 1; Length 610,
 Best Local Similarity 100.0%; Pred. No. 53,
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 LAGVIVA 39
 |||||
 DB 322 LAGVIVA 328

RESULT 22
 ID ABL_MOUSE STANDARD; PRT: 708 AA.
 AC Q9QXJ1: C08642;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Amyloid beta A4 precursor protein-binding family B member 1 (Fe65
 protein).
 GN ABL1 OR FE65.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Liakcheva A.V., Ivanova N.B., Belyavsky A.V.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 498-674 FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=97049965; PubMed=8894693;
 RA Bressler S.L., Gray M.D., Sopher B.L., Hu Q., Hearn M.G., Pham D.G.,
 RA Dinulos M.R., Pukroh K., Sisodia S.S., Miller M.A., Distchech C.M.,
 RA Martin G.M.;
 RT "cDNA cloning and chromosome mapping of the human Fe65 gene;
 RT interaction of the conserved cytoplasmic domains of the human beta-
 RT amyloid precursor protein and its homologues with the mouse Fe65
 RT protein.";
 RL Hum. Mol. Genet. 5:1589-1598(1996).
 CC -1- FUNCTION: MAY MODULATE THE INTERNALIZATION OF BETA-AMYLOID
 CC PEPCUSOR PROTEIN (BY SIMILARITY).
 CC -1- SUBUNIT: BINDS TO THE INTRACELLULAR DOMAIN OF THE BETA-AMYLOID
 CC PEPCUSOR PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 WW DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 PID DOMAINS.

CC CC This SWISS-Prot entry is copyright. It is produced through
 CC between the Swiss Institute of Bioinformatics and the
 CC the European Bioinformatics Institute. There are no fees
 CC use by non-profit institutions as long as its content
 CC modified and this statement is not removed. Usage by
 CC entities requires a license agreement (see <http://www.ebi.ac.uk>
 CC or send an email to license@sb-ebi.ch).
 DR EMBL: AE206720; AAF20141;
 DR EMBL: L77865; AAB51603.1;
 DR MGD: MGI:107765; Appb1.
 DR InterPro: IPR000050; PID_domain.
 DR Pfam: PF00397; WW; 1.
 DR Pfam: PF00640; PID; 2.
 DR SMART: SM00452; PTB; 2.
 DR SMART: SM00456; WW; 1.
 DR PROSITE: PS01179; PID; 2.
 DR PROSITE: PS01159; WW_DOMAIN_1; 1.
 DR PROSITE: PS50020; WW_DOMAIN_2; 1.
 KW Repeat.
 FT DOMAIN 253 285 WW.
 FT DOMAIN 379 507 PID 1.
 FT DOMAIN 540 697 PID 2.
 FT DOMAIN 158 171 GID-RICH.
 SV SEQUENCE 708 AA; 77138 MW; EFB978176E9898F2 CRC64;

Query Match 2.28; Score 7; DB 1; Length 708,
 Best Local Similarity 100.0%; Pred. No. 40,
 Matches 7; Conservative 0; Mismatches 0; Indels 0;

OY 47 VPSLSQ 54
 |||||
 DB 3 VPSLSQ 9

RESULT 24
 ID ABL_HUMAN STANDARD; PRT: 710 AA.
 AC Q00213;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Amyloid beta A4 precursor protein-binding family B mem-
 DE protein).
 GN ABL1 OR FE65.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eute-
 OC Mammalia; Eutheria; Primates; Caratheriina; Hominidae; Hu-
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Liakcheva A.V., Ivanova N.B., Belyavsky A.V.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 498-674 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97049965; PubMed=8894693;
 RA Bressler S.L., Gray M.D., Sopher B.L., Hu Q., Hearn M.G.,
 RA Dinulos M.R., Pukroh K., Sisodia S.S., Miller M.A.,
 RA Martin G.M.;
 RT "cDNA cloning and chromosome mapping of the human Fe65
 RT interaction of the conserved cytoplasmic domains of the
 RT amyloid precursor protein and its homologues with the hu-
 RT protein.";
 RL Hum. Mol. Genet. 5:1589-1598(1996).
 CC -1- FUNCTION: MAY MODULATE THE INTERNALIZATION OF BETA-AMYLOID
 CC PEPCUSOR PROTEIN (BY SIMILARITY).
 CC -1- SUBUNIT: BINDS TO THE INTRACELLULAR DOMAIN OF THE BETA-AMYLOID
 CC PEPCUSOR PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 WW DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 PID DOMAINS.


```

CC -1- FUNCTION: CATALYZES THE CROSS-LINKING OF PROTEINS AND THE
CC CONJUGATION OF POLYAMINES TO PROTEINS. CONTRIBUTES TO THE
CC FORMATION OF THE CORNIFIED CELL ENVELOPE OF KERATINOCYTES.
CC -1- CATALYTIC ACTIVITY: Protein glutamine + alkylamine = protein N(5)-
CC alkylglutamine + NH(3).
CC -1- COFACTOR: Binds 1 calcium ion per subunit (by similarity).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE)
CC AND A SHORT FORM. ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN FORESKIN KERATINOCYTES.
CC -1- SIMILARITY: BELONGS TO THE TRANSGLUTAMINASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF05960; AAC02978.1;
DR EMBL: AF05961; AAC02979.1;
DR EMBL: AF206510; AAF23981.1;
DR EMBL: AF206502; AAF23981.1; JOINED.
DR EMBL: AF206503; AAF23981.1; JOINED.
DR EMBL: AF206504; AAF23981.1; JOINED.
DR EMBL: AF206505; AAF23981.1; JOINED.
DR EMBL: AF206506; AAF23981.1; JOINED.
DR EMBL: AF206507; AAF23981.1; JOINED.
DR EMBL: AF206508; AAF23981.1; JOINED.
DR EMBL: AF206509; AAF23981.1; JOINED.
DR HSSP: P00488; IQK.
DR GENE: HGNC:11781; TGM5.
DR MIM: 603805.
DR InterPro: IPR001102; Glutaminase.
DR InterPro: IPR002931; Transglutaminase-like.
DR Pfam: PF00868; Transglutamin_N; 1.
DR Pfam: PF01841; Transglut-core; 2.
DR SMART: SM00460; TCG; 1.
DR PROSITE: PS00547; TRANSGLUTAMINASES; 1.
KW Transferrase; Acyltransferase; Calcium-binding; Polymorphism;
KW Alternative splicing.
FT ACT_SITE 278 278 BY SIMILARITY.
FT ACT_SITE 337 337 BY SIMILARITY.
FT ACT_SITE 360 360 BY SIMILARITY.
FT CA_BIND 400 400 BY SIMILARITY.
FT CA_BIND 402 402 BY SIMILARITY.
FT CA_BIND 448 448 BY SIMILARITY.
FT CA_BIND 453 453 BY SIMILARITY.
FT VARSPPLIC 64 145 MISSING (IN SHORT ISOFORM).
FT VARIANT 67 67 P -> S.
FT VARIANT 352 352 /FTID-VAR_013248.
FT VARIANT 352 352 A -> G.
FT SEQUENCE 720 AA: 80777 MW: 9658884.848kDa PRC64.
FT
Query Match 2.28; Score 7; DB 1; Length 720;
Best Local Similarity 100.0%; Pred No 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

DE 1/CR3/p150,95 beta-subunit) (title) (Complement receptor
DE subunit).
DE ITGB2 OR CD18.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Eute
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Bovidae
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RP SEQUENCE FROM N.A.
RX MEDLINE=930287; PubMed 1451021;
RA Shuster D.E., Bosworth R.T., Kehrl M.F., Jr.;
RT "Sequence of the bovine CD18-encoding cDNA: comparison w
RT and murine glycoproteins."
PL Gene 114:267-271(1992).
RN [2]
RP VARIANT LAD GLY-128.
RX MEDLINE=93028437; PubMed-1440465;
RA Shuster D.E., Kehrl M.F., Jr., Ackermann M.R., Gilbert P.
RT "Identification and prevalence of a genetic defect that i
RT leukocyte adhesion deficiency in Holstein cattle."
PL Proc Natl Acad Sci U S A 89:9245-9249(1992).
CC FUNCTION: INTEGRIN ALPHA-1/BEI-2 IS A RECEPTOR FOR
CC ICAM3 AND ICAM4. INTEGRINS ALPHA-M/BETA-2 AND ALPHA-X
CC AND FOR FIBRIN-GEN. INTEGRIN ALPHA-X/BETA-2 RECOGNIZE
CC SEQUENCE G-P-R IN FIBRIN-GEN ALPHA-CHAIN. INTEGRIN A
CC RECOGNIZES P1 AND P2 PEPTIDES OF FIBRINOGEN GAMMA CHA
CC ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR FACTOR X. INTE
CC D/BETA-2 IS A RECEPTOR FOR ICAM3 AND VCAM1.
CC -1- SHUNTIN. HEPHODIMER OF AN ALPHA AND A BETA SUBUNIT.
CC ASSOCIATES WITH EITHER AN ALPHA-L, ALPHA-M, ALPHA-X OR P
CC -1- SUBCELLULAR LOCATION: Type 1 membrane protein.
CC -1- DISEASE: DEFECTS IN ITGA2 ARE THE CAUSE OF LEUKOCTE
CC DEFICIENCY (LAD). THE MUTATION CAUSING LAD (GLY-128)
CC AMONG HOLSTEIN CATTLE THROUGHOUT THE WORLD, PLACING I
CC AMONG THE MOST COMMON GENETIC DISEASES KNOWN IN ANIMA
CC AGRICULTURE. ALL CATTLE WITH THE MUTANT ALLELE ARE PR
CC BULL. WHO THROUGH THE USE OF ARTIFICIAL INSEMINATION
CC CALVES IN THE 1950S AND 1960S.
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMIL
CC -1- SIMILARITY: CONTAINS 1 WFA-LIKE DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced throu
CC between the Swiss Institute of Bioinformatics and the
CC the European Bioinformatics Institute. There are no res
CC use by non-profit institutions as long as its content is
CC modified and this statement is not removed. Usage by a
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M81233; AAA30438.1;
DR FIC: J01121; J01121.
DR HSSP: P05106; J0V2.
DR InterPro: IPR000561; EGF-1-like.
DR InterPro: IPR002464; Integrin_B.
DR InterPro: IPR001169; Integrin_beta_C.
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR002035; VWF_A.
DR Pfam: PF00362; Integrin_B_1.
DR PRINTS: PR01186; INTEGRINB.
DR PRODOM: PD001811; Integrin_B_1.
DR SMART: SM00187; INB; 1.
DR SMART: SM00423; PST; 1.
DR SMART: SM00327; VWA; 1.
DR PROSITE: PS00243; INTEGRIN_BETA_3.
DR PROSITE: PS00022; EGF_1; EGF-WR-WR-2.
DR PROSITE: PS01186; EGF_2; UNKN-WR-2.
KW Integrin; Cell adhesion; Receptor; Transmembrane; Glycop
KW Repeat; Signal; Disease mutation.
FT CHAIN 1 22 BY SIMILARITY.
FT CHAIN 23 769 INTEGRIN_BETA-2.
FT DOMAIN 23 700 EXTRACELLULAR (POTENTIAL).

```

NO	SEQUENCE	7bp AA:	844bp MW:	500kDaP4E1406RFA C10064:
F1	TRANSMEM	701	723	POTENTIAL.
F1	DOMAIN	724	759	CYTOPLASMIC (POTENTIAL).
F1	DOMAIN	124	363	WFA-LIKE.
F1	REPEAT	449	617	4 CYSTEINE-RICH TANDEM REPEATS.
F1	REPEAT	449	496	1.
F1	REPEAT	497	540	11.
F1	REPEAT	541	581	111.
F1	REPEAT	582	617	TV
F1	DISULFID	25	447	BY SIMILARITY
F1	DISULFID	33	43	BY SIMILARITY.
F1	DISULFID	36	73	BY SIMILARITY.
F1	DISULFID	46	62	BY SIMILARITY.
F1	DISULFID	191	198	BY SIMILARITY.
F1	DISULFID	246	286	BY SIMILARITY.
F1	DISULFID	386	400	BY SIMILARITY.
F1	DISULFID	420	662	BY SIMILARITY.
F1	DISULFID	445	449	BY SIMILARITY.
F1	DISULFID	459	470	BY SIMILARITY.
F1	DISULFID	467	506	BY SIMILARITY.
F1	DISULFID	472	481	BY SIMILARITY
F1	DISULFID	483	497	BY SIMILARITY
F1	DISULFID	512	517	BY SIMILARITY.
F1	DISULFID	514	549	BY SIMILARITY.
F1	DISULFID	519	534	BY SIMILARITY.
F1	DISULFID	536	541	BY SIMILARITY
F1	DISULFID	557	562	BY SIMILARITY.
F1	DISULFID	559	590	BY SIMILARITY.
F1	DISULFID	564	573	BY SIMILARITY.
F1	DISULFID	575	582	BY SIMILARITY.
F1	DISULFID	596	601	BY SIMILARITY.
F1	DISULFID	598	643	BY SIMILARITY.
F1	DISULFID	603	612	BY SIMILARITY
F1	DISULFID	615	618	BY SIMILARITY.
F1	DISULFID	622	631	BY SIMILARITY.
F1	DISULFID	628	695	BY SIMILARITY.
F1	DISULFID	647	670	BY SIMILARITY.
F1	SITE	397	399	CELL ATTACHMENT SITE (POTENTIAL.)
F1	CARBOHYD	50	50	N-LINKED (GLCNAC. . .) (POTENTIAL.)
F1	CARBOHYD	116	116	N-LINKED (GLCNAC. . .) (POTENTIAL.)
F1	CARBOHYD	254	254	N-LINKED (GLCNAC. . .) (POTENTIAL.)
F1	CARBOHYD	501	501	N-LINKED (GLCNAC. . .) (POTENTIAL.)
F1	CARBOHYD	642	642	N-LINKED (GLCNAC. . .) (POTENTIAL.)
F1	VARIANT	128	128	D > G (IN LAD).
NO	SEQUENCE	7bp AA:	844bp MW:	500kDaP4E1406RFA C10064:

```

Query Match      2.2%  Score 7: 29.11  Length 769;
  Host Local Similarity 100.0%,  Prod. No. 65;
Matches      7:  Conservative      0:  Mismatches      0:  Indels      0:  Gaps
QY      71  AVEGLSE 77
      |||||
DB      441  AVEGLSE 347

RESULT 27
11B2_HUMAN
AC  P05107: 016418;          STANDARD;          PRI: 769 AA.
D1  13-AUG-1987 (Rel. 05, Created)
D7  01-APR-1988 (Rel. 07, Last sequence update)
D1  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Integrin beta-2 precursor (Cell surface adhesion glycoprotein LFA-
DE  1/2P/CD10.35 beta subunit) (CD18) (Complement receptor 2 beta
DE  subunit).
DE  Homo sapiens (Human).
GN  H102 OR CD18.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX  NCBI-TaxID=9606;
RN  11
RP  SEQUENCE FROM N.A.
RX  MEDLINE 87131980; PubMed 30286466;
RA  Kishimoto T K., O'Connor K., Lee A., Eckerts T.M., Springer T.A.,

```

RA "Cloning of the beta subunit of the leukocyte adhesion proteins:
 biology of an extracellular matrix receptor defines a novel
 RT supergene family." ;
 RL Cell 48:1681-690(1987).
 RN
 RN SEQUENCE FROM N.A.
 RP MEDLINE=92077153; PubMed=1684838;
 RX Weinstein J.R., Wells C.E., Wright A.H., Clark P.A., Law S.K.A.;
 FT "The gene organisation of the human beta 2 integrin subunit (CD18).";
 RL FEMS Lett. 294:97-103(1991).
 RN
 RN SEQUENCE FROM N.A.
 RP MEDLINE=92089799; PubMed=10840953;
 RX Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
 PA Park H.-S., Toyoda A., Ishii K., Tokiaki Y., Choi B.-K., Soda E.,
 GA Giki M., Takaki T., Sakaki Y., Tawada S., Hieshimauchi K., Polley A.,
 PA Menzel U., Polahar J., Knopf K., Lehmann P., Patterson D.,
 PA Persechini K., Rupp T., Shihara K., Kawasaki K., Asakawa S.,
 PA Shintani A., Sasaki T., Nayamine K., Mitsuyama S., Antizaralis S.E.,
 PA Shostakov S., Shemitz N., Nayamine K., Horstischer K., Brandt P.,
 PA Sharif M., Sison G., Desario A., Reibelt J., Kang G., Blocker H.,
 PA Ramser J., Brock A., Klages S., Henning S., Kieselmann L., Nagend F.,
 PA Weinmeyer S., Bergem K., Girlnot K., Nitzel D., Francis F.,
 PA Lehmann B., Reinhardt P., Yipso M.-E.;
 FT "The DNA sequence of human chromosome 21." ;
 RL Nature 405:311-319(2000).
 RN
 RN SEQUENCE OF 9-769 FROM N.A. AND PARTIAL SEQUENCE.
 RP TISSUE-Spleen;
 RX MEDLINE=92346525; PubMed=2954816;
 PA Law S.K.A., Garrison J., Hiltzsch J.E., Wells C.E., Willis A.C.,
 RA Wong A.U.;
 FT "The primary structure of the beta-subunit of the cell surface
 RT adhesion glycoproteins LFA-1, CD3 and p150.95 and its relationship to
 RL the fibronectin receptor." ;
 RL PMOL 1: 6-915-919(1987).
 RN
 RN SEQUENCE OF 447-355 FROM N.A. AND VARIANTS LAD1 SEP-351 AND IR0-586.
 RP MEDLINE=92170641; PubMed=134613;
 RX Neilson G., Rubb H., Arnaout M.A.;
 FT "Genetic cause of leukocyte adhesion molecule deficiency: Altered
 RT splicing and a missense mutation in a conserved region of CD18 impair
 RL cell surface expression of beta 2 integrins." ;
 RL J. Biol. Chem. 267:3351-3357(1992).
 RN
 RN VARIANTS LAD1 PRO-149 AND ARG-159.
 RP MEDLINE=9024666; PubMed=1694220;
 RX Wardlaw A.J., Hibbs M.L., Stoecker S.A., Springer T.A.;
 RA "Distal mutations in two patients with leukocyte adhesion
 RT deficiency and their functional correlates." ;
 RL J. Exp. Med. 172:335-345(1990).
 RN
 RN VARIANT LAD1 ARG-159.
 RP MEDLINE=92134403; PubMed=1352501;
 RX Corbi A., Vera A., Dpsa A., Rodriguez M.C.G., Fontan G.,
 RA Sanchez-Madrid F.;
 FT "Molecular basis for a severe case of leukocyte adhesion deficiency." ;
 RL Eur. J. Immunol. 22:1877-1881(1992).
 RN
 RN VARIANTS LAD1 THR-106 AND CYS-593.
 RP MEDLINE=90187099; PubMed=1966941;
 RX Attard M.A., and N., Gupta R., Leach G.G., Fathallah D.M.;
 FT "Point mutations impairing cell surface expression of the common beta
 RT subunit (CD18) in a patient with leukocyte adhesion molecule
 RL (Leu-CAM) deficiency." ;
 RL J. Clin. Invest. 85:977-981(1990).
 RN
 RN VARIANT LAD1 LEU-178.
 RP MEDLINE=92184805; PubMed=1347532;
 RX Back L.L., Kwok W.W., Hiestein D.D.;
 FT "Identification of two molecular defects in a child with leukocyte
 RT adhesion deficiency." ;

PI J Biol Chem. 267:5482-5487(1992).

RN [10]

RP VARIANT LAD1 ASN-128.

RX MEDLINE-92272746; PubMed-1590804.

RA Matsura S., Kishi F., Tsukahara M., Nunoi H., Matsuda I.,

KA Kobayashi K., Kajii T.

RT "Leukocyte adhesion deficiency: identification of novel mutations in

KL two Japanese patients with a severe form."

RN [11]

RP VARIANT LAD1 SER-284.

RX MEDLINE-9312347; PubMed-7686755;

KA Back L.A., Kerkering M., Baker D., Hauser J.R., Embrey L.J.,

KA Hinkley D.D.

RT "A point mutation associated with leukocyte adhesion deficiency type

KL 1 of moderate severity."

RN [12]

RP Biochem Biophys Res Commun 193:912-918(1993)

RP VARIANT LAD1 PRO-138 AND ARG-273.

RX MEDLINE-94102253; PubMed-9884334;

KA Hogg N., Stewart M.P., Smith S.L., Newton P., Shaw T.M., Jew S.K.A.,

KA Klein N.

RT "A novel leukocyte adhesion deficiency caused by expressed but

KL nonfunctional beta2 integrins Mac-1 and LFA-1."

RN [13]

RP J Clin Invest. 103:97-106(1999).

RT "FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2,

CC ICAM3 AND ICAM4. INTEGRINS ALPHA-M/BETA-2 AND ALPHA-X/BETA-2 ARE

CC RECEPTORS FOR THE IC3H FRAGMENT OF THE THIRD COMPLEMENT COMPONENT

CC AND FOR FIBRINOGEN. INTEGRIN ALPHA-X/BETA-2 RECOGNIZES THE

CC SEQUENCE G-P-R IN FIBRINOGEN ALPHA-CHAIN. INTEGRIN ALPHA-M/BETA-2

CC RECOGNIZES P1 AND P2 PEPTIDES OF FIBRINOGEN GAMMA CHAIN. INTEGRIN

CC ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR FACTOR X. INTEGRIN ALPHA-

CC D/BETA-2 IS A RECEPTOR FOR ICAM3 AND VCAM1.

CC SUBUNIT HETEROPTIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-2

CC ASSOCIATES WITH EITHER ALPHA-L, ALPHA-M, ALPHA-X OR ALPHA-D.

CC -1- SUBCELLULAR LOCATION: Type 1 membrane protein.

CC -1- DISEASE: DEFECTS IN ITGB2 ARE THE CAUSE OF LEUKOCYTE ADHESION

CC DEFICIENCY, TYPE 1 (LAD1). LAD1 PATIENTS HAVE RECURRENT BACTERIAL

CC INFECTIONS AND THEIR LEUKOCYTES ARE DEFICIENT IN A WIDE RANGE OF

CC ADHESION-DEPENDENT FUNCTIONS.

CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.

CC -1- SIMILARITY: CONTAINS 1 VWA-LIKE DOMAIN.

CC -1- DATABASE: NAME-PROV: NOTE-CD guide cd18 entry;

CC WWW-http://www.ncbi.nlm.nih.gov/prov/cd/cd18.htm".

CC -----

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation-

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC -----

DR EMBL: Y00057; CAA68266.1; -

DR EMBL: X64072; CAA45427.1; -

DR EMBL: X64073; CAA45427.1; JOINED.

DR EMBL: X64074; CAA45427.1; JOINED.

DR EMBL: X64075; CAA45427.1; JOINED.

DR EMBL: X64076; CAA45427.1; JOINED.

DR EMBL: X64077; CAA45427.1; JOINED.

DR EMBL: X64078; CAA45427.1; JOINED.

DR EMBL: X64079; CAA45427.1; JOINED.

DR EMBL: X64080; CAA45427.1; JOINED.

DR EMBL: X64081; CAA45427.1; JOINED.

DR EMBL: X64082; CAA45427.1; JOINED.

DR EMBL: X64083; CAA45427.1; JOINED.

DR EMBL: X64924; CAA45427.1; JOINED.

DR EMBL: X64925; CAA45427.1; JOINED.

DR EMBL: X64926; CAA45427.1; JOINED.

DR EMBL: AL163300; CAB90553.1; -

DR EMBL: M15195; AAA59490.1; -

DR EMBL: S81234; AAB21404.1; -

DR PIR: A25967; LJDHLM.

DR PIR: A25965; A25965.

DR HSSP: P05106; 1JV2.

DR Genew: HGNC:6155; ITGB2.

DR MIM: 600085; -

DR MIM: 116920; -

DR InterPro: IPR000641; PPF-1 like.

DR InterPro: IPR002869; Integrin_b.

DR InterPro: IPR001694; Integrin_beta_2.

DR InterPro: IPR003659; Plexin-like.

DR InterPro: IPR002035; VWF_A.

DR Pfam: PF00362; Integrin_B_1.

DR PRINTS: PD01186; INTEGRIN.

DR PRODOM: PD01811; Integrin_b_1.

DR SMART: SM00187; INH_1.

DR SMART: SM00424; PS1_1.

DR SMART: SM00327; VWA_1.

DR PROSITE: PS00243; INTEGRIN_BETA_2.

DR PROSITE: PS00222; EGF_1; CRK-AML_2.

DR PROSITE: PS01186; EGF_2; UNKNOWN_4.

DR Integrin: Cell adhesion, Receptor, Transmembrane; Glycyl

DR Repeat, Signal, Disease mutation.

DR SIGNAL 1 22

FT SIGNAL 1 22

FT CHAIN 23 769

FT DOMAIN 23 700

FT TRANSMEM 701 723

FT DOMAIN 724 769

FT DOMAIN 124 363

FT DOMAIN 449 617

FT REPEAT 449 496

FT REPEAT 497 540

FT REPEAT 541 581

FT REPEAT 582 617

FT DISULFID 25 447

FT DISULFID 33 43

FT DISULFID 36 73

FT DISULFID 46 62

FT DISULFID 191 198

Query Match 2.28; Score 7; DB: 1; Length 75

Best Local Similarity 100.0%; Prod. No: 69;

Matches 7; Conservative 0; Mismatches 0

QY 71 ANGELSE 77

DB 341 ANGELSE 347

RESULT 28

ITB2_PIG 1162_PIG STANDARD; EPI: 769 AA.

AC P53714;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Integrin beta-2 precursor (ccl) surface adhesion glycoprotein

DE I/CR3/β156,95 beta-subunit) (CluB) (Complement receptor

DE subunit).

GN ITGB2 OR CD18.

OS Sus scrofa (Pig).

OC Mammalia, Metazoa, Chordata, Vertebrata, Art.

OC Mammalia, Eutheria, Cetartiodactyla, Suidae, Sus

OX NCBI_TaxID=9823;

ON [1]

RP SEQUENCE FROM N.A.

RA Lee J.K., Strick L.B., Rothbard M.S.

RT "Molecular cloning and characterization of the porcine

RT Acetabularia sp. 3:222-240(1996).

FL -1- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR

CC ICAM3 AND ICAM4. INTEGRINS ALPHA-M/BETA-2 AND ALPHA-

CC RECEPTORS FOR THE IC3H FRAGMENT OF THE THIRD COMPLEMENT

CC AND FOR FIBRINOGEN. INTEGRIN ALPHA-X/BETA-2 RECOGN-

CC SEQUENCE G-P-R IN FIBRINOGEN ALPHA-CHAIN. INTEGRIN

PROGAMIZES P1 AND P2 PEPTIDES OF FIBRINOGEN GAMMA CHAIN. INTERFERIN
 CC ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR FACTOR X. INTEGRIN ALPHA-
 CC D/BETA-2 IS A RECEPTOR FOR TROMB AND VCAM1.
 CC -1- SRRNIT. HETEROFORMER OF AN ALPHA AND A BETA SRRNIT. BETA-2
 CC ASSOCIATES WITH EITHER ALPHA-L, ALPHA-M, ALPHA-X OR ALPHA-D.
 CC -1- SUPRACELLULAR LOCATION. TYPE 1 MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 VWFA-LIKE DOMAIN.
 CC -----
 CC THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
 CC BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -
 CC THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
 CC USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
 CC MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
 CC ENTITIES REQUIRES A LICENSE AGREEMENT (SEE <http://www.ist.sib.ch/unicomney>
 CC OR SEND AN EMAIL TO license@sib.ch).
 CC -----
 CC FMH01: 014941; AAB16868.1; -
 CC HSS01: P05106; JUV2.
 CC INTERPRO: IPR000561; EGF_LIKE.
 CC INTERPRO: IPR002389; Integrin_B.
 CC INTERPRO: IPR001169; Integrin_beta_C.
 CC INTERPRO: IPR003659; Plexin_LIKE.
 CC INTERPRO: IPR002835; VWFA.
 CC Pfam: PF00362; Integrin_B; 1.
 CC PRINTS: PK01186; INTEGRINB.
 CC PRODOM: PD001811; Integrin_B; 1.
 CC SMART: SM00423; PSI; 1.
 CC SMART: SM00327; VWFA; 1.
 CC PROSITE: PS00243; INTEGRIN_BETA_3.
 CC PROSITE: PS00622; EGF_2; UNKNOWN_3.
 CC PROSITE: PS01186; EGF_2; UNKNOWN_3.
 CC INTERPRO: Cell adhesion; Transmembrane; Glycophorin; Fc-gamma; Signal
 CC FT CHAIN 1 22
 CC FT DOMAIN 23 769 INTEGRIN BETA-2
 CC FT TRANSMEM 761 728 EXTRACELLULAR (POTENTIAL).
 CC FT DOMAIN 724 769 POTENTIAL.
 CC FT DOMAIN 124 363 CYTOPLASMIC (POTENTIAL).
 CC FT DOMAIN 449 617 VWFA-LIKE.
 CC FT REPEAT 449 496 4 CYSTEINE-RICH TANDEM REPEATS.
 CC FT REPEAT 497 540 1.
 CC FT REPEAT 541 581 11.
 CC FT REPEAT 582 617 11.
 CC FT REPEAT 617 647 11.
 CC FT DISULFID 25 447 IV.
 CC FT DISULFID 33 43 BY SIMILARITY.
 CC FT DISULFID 36 73 BY SIMILARITY.
 CC FT DISULFID 46 62 BY SIMILARITY.
 CC FT DISULFID 191 198 BY SIMILARITY.
 CC FT DISULFID 246 286 BY SIMILARITY.
 CC FT DISULFID 386 400 BY SIMILARITY.
 CC FT DISULFID 420 662 BY SIMILARITY.
 CC FT DISULFID 445 449 BY SIMILARITY.
 CC FT DISULFID 459 470 BY SIMILARITY.
 CC FT DISULFID 467 506 BY SIMILARITY.
 CC FT DISULFID 472 481 BY SIMILARITY.
 CC FT DISULFID 483 497 BY SIMILARITY.
 CC FT DISULFID 512 517 BY SIMILARITY.
 CC FT DISULFID 514 549 BY SIMILARITY.
 CC FT DISULFID 519 534 BY SIMILARITY.
 CC FT DISULFID 536 541 BY SIMILARITY.
 CC FT DISULFID 541 562 BY SIMILARITY.
 CC FT DISULFID 557 562 BY SIMILARITY.
 CC FT DISULFID 559 590 BY SIMILARITY.
 CC FT DISULFID 564 573 BY SIMILARITY.
 CC FT DISULFID 575 582 BY SIMILARITY.
 CC FT DISULFID 596 601 BY SIMILARITY.
 CC FT DISULFID 598 643 BY SIMILARITY.
 CC FT DISULFID 603 612 BY SIMILARITY.
 CC FT DISULFID 615 618 BY SIMILARITY.
 CC FT DISULFID 622 631 BY SIMILARITY.
 CC FT DISULFID 628 695 BY SIMILARITY.
 CC FT DISULFID 647 670 BY SIMILARITY.

FT SITE 397 399 CELL ATTACHMENT SITE (POTENTIAL).
 FT CAPRHQHD 50 50 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CAPRHQHD 116 116 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CAPRHQHD 254 254 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CAPRHQHD 501 501 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CAPRHQHD 642 642 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 769 AA: 84789 MW: 19066.665 KDa
 Best Match 2.2% Score 71 DB 1; Length 769;
 Query Local Similarity 100.0% Pct-Id No. 65;
 Matches 71 Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 71 ANCHOR 77
 1th 341 ANCHOR 347
 RESULT 29
 ID 1TB7_HUMAN STANAPRO: PRT: 798 AA
 AC P26010;
 DT 01-MAY-1992 (Ref. 22, Created)
 DT 01-MAY-1992 (Ref. 22, Last sequence update)
 DT 15-JUN-2002 (Ref. 41, Last annotation update)
 DE Integrin beta-7 precursor.
 GN ITGB7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catartia; Homiidae; Homo.
 OX NCBI_TaxID: 9606;
 RN [1]
 RP SEQUENCE FROM N.A. (SHORE AND LONG) (PDBMS).
 RC TISSUE: Leukocyte.
 FX REFSEQ: NM_004065, PubMed: 2040616;
 PA Pile D. T. Purogic C., Sheppard P., Pytela P.;
 KT "Complete amino acid sequence of an integrin beta subunit (beta 7)
 identified in leukocytes.";
 PL J Biol Chem 266:11099-11016(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE: 91190778; PubMed: 2083230;
 RX Yuan H., Jiang W.-M., Kristensen G.W., Watson J.D.;
 KT "Cloning and sequence analysis of a novel beta 2-related integrin
 transcript from T lymphocytes: homology of integrin cysteine-rich
 repeats to domain III of laminin B chains.";
 RL Int. Immunol. 2:1097-1108(1990).
 RN [3]
 RP REVISIONS.
 FX MEDLINE: 92135083; PubMed: 1777426;
 RX Yuan H., Jiang W.-M., Kristensen G.W., Watson J.D.;
 KT "Cloning and sequence analysis of a novel beta 2-related integrin
 transcript from T lymphocytes: homology of integrin cysteine-rich
 repeats to domain III of laminin B chains.";
 RL Int. Immunol. 3:1373-1374(1991).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 93002753; PubMed: 1382574;
 RX Jiang W.-M., Jenkins D., Yuan Q., Leung E., Choo K.H., Watson J.D.,
 RA Kristensen G.W.;
 FT "The gene organization of the human beta 7 subunit, the common beta
 subunit of the leukocyte integrins HML-1 and LPAM-1.";
 RL Int. Immunol. 4:1031-1040(1992).
 RN [5]
 RP MOTAGNESIS OF ASP-159.
 RX MEDLINE: 20400502; PubMed: 10837471;
 RX Higgins J.M.G., Bernadas M., Tan K., Irie A., Wang J.-H., Takada Y.,
 RA Brenner M.B.;
 FT "The role of alpha and beta chains in ligand recognition by beta 7
 integrins.";
 RL J Biol Chem. 275:25652-25664(2000).
 CC -1- FUNCTION: INTEGRIN ALPHA-4/BETA-7 (PEYER'S PATCHES-SPECIFIC HOMING
 RECEPTOR LPAM-1) IS EXPECTED TO PLAY A ROLE IN ADHESIVE
 INTERACTIONS OF LEUKOCYTES. IT IS A RECEPTOR FOR FIBRONECTIN AND

CC RECOGNIZES ONE OR MORE DOMAINS WITHIN THE ALTERNATIVELY SPLICED
 CC CS-1 REGION OF FIBRONECTIN. INTEGRIN ALPHA-4/BETA-7 IS ALSO A
 CC RECEPTOR FOR MADCAM1 AND VCAM1. IT RECOGNIZES THE SEQUENCE L-D-T
 CC IN MADCAM1. INTEGRIN ALPHA-E/BETA-7 (HML-1) IS A RECEPTOR FOR E-
 CC CADHERIN.
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-7
 CC ASSOCIATES WITH EITHER ALPHA-4 OR ALPHA-E.
 CC -1- SUBCELLULAR LOCATION: Type 1 membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM. ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF LEUKOCYTE LINES.
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 VWA-LIKE DOMAIN.
 CC
 CC THIS SWISS-PROT entry is copyright © It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation at
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a separate agreement. (See <http://www.isb-sib.ch/announcements/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: S80335; AAB21332.1; -;
 DR EMBL: M62880; AAA59185.1; -;
 DR EMBL: M68892; AAA59184.1; -;
 DR EMBL: S49378; AAB21688.1; -;
 DR EMBL: S49364; AAB21688.1; JOINED.
 DR EMBL: S49365; AAB21688.1; JOINED.
 DR EMBL: S49366; AAB21688.1; JOINED.
 DR EMBL: S49367; AAB21688.1; JOINED.
 DR EMBL: S49368; AAB21688.1; JOINED.
 DR EMBL: S49369; AAB21688.1; JOINED.
 DR EMBL: S49370; AAB21688.1; JOINED.
 DR EMBL: S49371; AAB21688.1; JOINED.
 DR EMBL: S49373; AAB21688.1; JOINED.
 DR EMBL: S49374; AAB21688.1; JOINED.
 DR EMBL: S49375; AAB21688.1; JOINED.
 DR EMBL: S49377; AAB21688.1; JOINED.
 DR EMBL: L23823; AAA36118.1; -;
 DR EMBL: L23810; AAA36118.1; JOINED.
 DR EMBL: L23811; AAA36118.1; JOINED.
 DR EMBL: L23812; AAA36118.1; JOINED.
 DR EMBL: L23813; AAA36118.1; JOINED.
 DR EMBL: L23814; AAA36118.1; JOINED.
 DR EMBL: L23815; AAA36118.1; JOINED.
 DR EMBL: L23816; AAA36118.1; JOINED.
 DR EMBL: L23817; AAA36118.1; JOINED.
 DR EMBL: L23818; AAA36118.1; JOINED.
 DR EMBL: L23819; AAA36118.1; JOINED.
 DR EMBL: L23820; AAA36118.1; JOINED.
 DR EMBL: L23822; AAA36118.1; JOINED.
 DR PIR: A40526; A40526.
 DR HSSP: P05106; IJVE.
 DR Gene: HGNC:6162; ITGB7.
 DR MIM: 147559; -;
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002369; Integrin_B.
 DR InterPro: IPR001169; Integrin_beta_C.
 DR InterPro: IPR003659; Plexin-like.
 DR InterPro: IPR002035; VWF_A.
 DR Pfam: PF00362; Integrin_B; 1.
 DR PRINTS: PR01186; INTEGRINB.
 DR ProDom: PD001811; Integrin_B; 1.
 DR SMART: SM00001; EGF_Like; 1.
 DR SMART: SM00187; INH; 1.
 DR SMART: SM00423; PSI; 1.
 DR PROSITE: PS00327; VWA; 1.
 DR PROSITE: PS00243; INTEGRIN_BETA_3.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_4.
 DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
 KW Integrin: Cell adhesion; Receptor; Transmembrane; Glycoprotein;
 KW Repeat: Signal; Phosphorylation; Alternative splicing.
 FT SIGNAL 1 19

FT	CHAIN	20	798	INTEGRIN BETA-7.
FT	DOMAIN	20	723	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	724	746	POTENTIAL.
FT	DOMAIN	747	798	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	150	389	VWA-LIKE.
FT	DOMAIN	478	640	4 CYSTEINE-RICH TANDEN REPE
FT	REPEAT	478	526	1.
FT	REPEAT	527	565	11.
FT	REPEAT	566	604	111.
FT	REPEAT	605	640	11.
FT	DISULFID	45	476	BY SIMILARITY.
FT	DISULFID	51	61	BY SIMILARITY.
FT	DISULFID	54	91	BY SIMILARITY.
FT	DISULFID	64	80	BY SIMILARITY.
FT	DISULFID	216	223	BY SIMILARITY.
FT	DISULFID	271	311	BY SIMILARITY.
FT	DISULFID	412	428	BY SIMILARITY.
FT	DISULFID	448	688	BY SIMILARITY.
FT	DISULFID	474	478	BY SIMILARITY.
FT	DISULFID	488	500	BY SIMILARITY.
FT	DISULFID	497	537	BY SIMILARITY.
FT	DISULFID	502	511	BY SIMILARITY.
FT	DISULFID	513	527	BY SIMILARITY.
FT	DISULFID	543	548	BY SIMILARITY.
FT	DISULFID	545	574	BY SIMILARITY.
FT	DISULFID	550	559	BY SIMILARITY.
FT	DISULFID	561	566	BY SIMILARITY.
FT	DISULFID	580	585	BY SIMILARITY.
FT	DISULFID	582	613	BY SIMILARITY.
FT	DISULFID	587	596	BY SIMILARITY.
FT	DISULFID	598	605	BY SIMILARITY.
FT	DISULFID	619	624	BY SIMILARITY.
FT	DISULFID	621	666	BY SIMILARITY.
FT	DISULFID	626	635	BY SIMILARITY.
FT	DISULFID	638	641	BY SIMILARITY.
FT	DISULFID	645	654	BY SIMILARITY.
FT	CARBOHYD	279	279	N-LINKED (GLCNAC...) (P...)
FT	CARBOHYD	434	434	N-LINKED (GLCNAC...) (P...)
FT	CARBOHYD	477	477	N-LINKED (GLCNAC...) (P...)
FT	CARBOHYD	531	531	N-LINKED (GLCNAC...) (P...)
FT	CARBOHYD	590	590	N-LINKED (GLCNAC...) (P...)
FT	CARBOHYD	665	665	N-LINKED (GLCNAC...) (P...)
FT	CARBOHYD	674	674	N-LINKED (GLCNAC...) (P...)
FT	MOD_RES	778	778	PHOSPHORYLATION (BY TYR-KIN...)
FT	VARSPLIC	501	648	MISSING (IN SHORT ISOFORM)
FT	MUTAGEN	159	159	D-EA: LOSS OF INTEGRIN ALI...
FT				BINDING TO E-CADHERIN AND
FT				ALPHA-4/BETA-7 BINDING TO E
FT	SEQUENCE	798 AA; 8650+ MW; 1982750E9992785 CR64;		

Query Match 2.2%; Score 71; DB 1; Length 798
 Best Local Similarity 100.0%; Prod. No. 67;
 Matches 7; Conservative 0; Mismatches 0; Indels 0;

Oy 71 AVGELSE 77
 Db 367 AVGLISE 374

RESULT 30
 SCYL_YEAST
 ID SCYL_YEAST STANDARD; (P): 804 AA
 AC P51009;
 DT 01-OCT-1996 (rel. 34, Cited)
 DT 01-OCT-1996 (rel. 34, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE SCYL protein.
 DE SCYL OR YGL083W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes.
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

```

CC NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
KA Petersen M., Gallwitz D.;
RL Submitted (Apr-1996) to the EMBL/GenBank/DDBJ databases
RN [2]
RP SEQUENCE FROM N.A.
KA STRAIN=5286;
RX MEDLINE=97435481; PubMed 9290212;
KA Rieper M., Riesecker M., Schaper M., Mueller-Auer S.;
RI Sequence analysis of 203 kinases from Saccharomyces cerevisiae
RT Chromosome VII."
RL Yeast 13:1077-1090(1997)
CC 1- SIMILARITY: SOME TO THE CATALYTIC DOMAIN OF PROTEIN KINASES
CC 2 This SWISS-Prot entry is copyright. It is produced through a collaboration
CC 3 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC 4 the European Bioinformatics Institute. There are no restrictions on its
CC 5 use by non-profit institutions as long as its content is in no way
CC 6 modified and this statement is not removed. Usage by and for commercial
CC 7 entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC 8 or send an email to license@sib.ch)
CC 9
D3 EMBL: X97305; CAA65975.1;
D3 EMBL: Z72605; CAA96788.1;
D3 SDD: S0003051; SCY1
D3 InterPro: IPR000719; Euk.pkinase.
D3 Pfam: PF000069; Pkinase.1
D3 PROSITE: PS50011; PROTEIN_KINASE_DOM.1
D3 DOMAIN 1 324 PROTEIN KINASE-LIKE.
D3 F 16 43 POLY-SER.
D3 SEQUENCE 804 AA; 91000 MW; 457A81DD7AB808E CRC64;
D3
Query Match 2.28; Score 7; DB 1; Length 804;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
D3 236 SEQFID 242
D3 134 ECONFID 140
D3
RESULT 31
YBBP_ECOLI
ID YBBP_ECOLI STANDARD; PRT; 804 AA.
AC P77504;
D1 01-NOV-1997 (Rel. 35, Created)
D1 01-NOV-1997 (Rel. 35, Last sequence update)
D1 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ybbp.
GN YBBP OF H0406.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
D3 [1]
RP SEQUENCE FROM N.A.
KA STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed 9278503;
KA Blattner F.R., Plunkett G. III, Bloch C.A., Bernal N.T., Burland V.,
KA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.W.,
KA Glover J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose T.J.,
KA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RI Science 277:1453-1474(1997).
D3 [2]
RP SEQUENCE FROM N.A.
KA Roberts R., Allen E., Araujo S., Aparicio A., Chung E., Davis K.,
KA Duran M., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi D.,
KA Lew H., Lin D., Nemach A., Oefner P., Schramm S., Davis P.W.;
KA Submitted (Jan-1997) to the EMBL/GenBank/DDBJ databases.
D3 1- SURPLICULAR LOCATION: Integral membrane protein (Potential);

```

```

CC CC
CC This SWISS-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib.ch)
CC 9
D3 EMBL: A6000155; AAC73598.1;
D3 EMBL: 082544; AAE40250.1;
D3 EcoGene: EG03263; ybbp.
D3 InterPro: IPR004838; DUF214.
D3 Pfam: PF02687; DUF214; 1.
D3 Hypothetical protein; Transmembrane; Complete proteome.
D3 FT TRANSMEM 15 35 POTENTIAL.
D3 FT TRANSMEM 24 263 POTENTIAL.
D3 FT TRANSMEM 301 321 POTENTIAL.
D3 FT TRANSMEM 334 353 POTENTIAL.
D3 FT TRANSMEM 381 401 POTENTIAL.
D3 FT TRANSMEM 403 423 POTENTIAL.
D3 FT TRANSMEM 454 473 POTENTIAL.
D3 FT TRANSMEM 680 700 POTENTIAL.
D3 FT TRANSMEM 734 754 POTENTIAL.
D3 FT TRANSMEM 769 789 POTENTIAL.
D3 SEQUENCE 804 AA; 89332 MW; 3475814069207A CRC64;
D3
Query Match 2.28; Score 7; DB 1; Length 804;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
D3 28 ISFMIL 34
D3 462 ISFMIL 468
D3
RESULT 32
ITB7_MOUSE
ID ITB7_MOUSE STAN/ABU; PRI; 806 AA.
AC P26011; O64656;
D1 01-MAY-1992 (Rel. 22, Created)
D1 01-JUL-1993 (Rel. 26, Last sequence update)
D1 15-JUN-2002 (Rel. 41, Last annotation update)
DE Integrin beta-7 precursor (Integrin beta-P) (M290 IEL antigen).
GN ITB7.
OS Mus musculus (Mouse).
OC Eukaryota; Mammalia; Chordata; Craciata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
D3 [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92390425; PubMed 151854;
KA Hu M.C.L., Brown P.T., Weissman I.L., Holzmann B.;
RT "Cloning and expression of mouse integrin beta p(bu7): a
RT functional role in Peyer's patch-specific lymphocyte homing."
RI Proc. Natl. Acad. Sci. U.S.A. 89:8754-8758(1992).
D3 [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92218384; PubMed 1579978;
KA Yuan Q., Jiang W. M., Leung K., Hollander D., Watson J.D.,
KA Kristiansen G.W.;
RT "Molecular cloning of the mouse integrin beta 7 subunit."
RI J. Biol. Chem. 267:7350-7356(1992).
D3 [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92488644; PubMed 138190;
KA Gurish M.F., Bell A.F., Smith T.J., Ducharme L.A., Wang R.K.,
KA Weis J.H.;
RT "Expression of murine beta 7, alpha 4, and beta 1 integrin genes by
RT rodent mast cells."
RI J. Immunol. 149:1964-1972(1992).
D3 [4]
RP SEQUENCE OF 1-50 FROM N.A. AND SEQUENCE OF 20-32.

```


RX MEDLINE:91248239, PubMed:1710115,
 RA Yuan Q., Jlang W.-M., Hollander D., Leung E., Watson J.D.,
 RA Kriksen G.W.,
 RT "Identity between the novel integrin beta 7 subunit and an antigen
 RT found highly expressed on intra-epithelial lymphocytes in the small
 RT intestine."
 RL Biochem. Biophys. Res. Commun. 176:1443-1449(1991).
 RN [5]
 RP SEQUENCE OF 1-67 FROM N.A.
 RA MEDLINE:93305607; PubMed:8318458;
 RA Leung E., Mead P.E., Yuan Q., Jlang W.M., Watson J.D.,
 RA Kriksen G.W.,
 RT "The mouse beta 7 integrin gene promoter: transcriptional regulation
 RT of the leukocyte integrins LPA-1 and M290."
 RL Immunol. 5:551-558(1993).
 CC -1- FUNCTION: INTEGRIN ALPHA-4/BETA-7 (PEYER'S PATCHES SPECIFIC HOMING
 CC RECEPTOR LPA-1) IS INVOLVED IN ADHESIVE INTERACTIONS OF
 CC LEUCOCYTES. IT IS A RECEPTOR FOR FIBRONECTIN AND RECOGNIZES ONE OR
 CC MORE DOMAINS WITHIN THE ALTERNATIVELY SPLICED CS-1 REGION OF
 CC FIBRONECTIN. INTEGRIN ALPHA-4/BETA-7 IS ALSO A RECEPTOR FOR
 CC MADCAM1 AND VCAM1. IT RECOGNIZES THE SEQUENCE L-D-T IN MADCAM1.
 CC INTEGRIN ALPHA-4/BETA-7 IS A RECEPTOR FOR E-CADHERIN.
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-7
 CC ASSOCIATES WITH EITHER ALPHA-4 OR ALPHA-E.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 WFPA-LIKE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed, usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M68903, AAA02749.1,
 DR EMBL: M95632, AAA39323.1,
 DR EMBL: M95633, AAA39324.1,
 DR EMBL: S44607, AAB23193.1,
 DR EMBL: S64504, AAB27396.1,
 DR PIR: P00017, P00017,
 DR PIR: A46271, A46271,
 DR PIR: A42483, A42483,
 DR HSP: P05106, IJ2,
 DR MGI: MGI:96616, Itgb7,
 DR InterPro: IPR000561, EGF-like,
 DR InterPro: IPR002369, Integrin_B,
 DR InterPro: IPR001169, Integrin_Beta_7,
 DR InterPro: IPR001659, Plexin-like,
 DR InterPro: IPR002035, WFPA-like,
 DR Pfam: PF00362, Integrin_B, 1,
 DR PRINTS: PR01186, INTEGRIN,
 DR PRODOM: PD001811, Integrin_B, 1,
 DR SMART: SM00187, INB, 1,
 DR SMART: SM00423, PSI, 1,
 DR SMART: SM00327, VWA, 1,
 DR PROSITE: PS00243, INTEGRIN_BETA_3,
 DR PROSITE: PS00022, EGF_1, UNKNOWN_4,
 DR PROSITE: PS01166, EGF_2, UNKNOWN_1,
 KM Integrin: Cell adhesion; Receptor; Transmembrane; Glycoprotein;
 FT Repeat: Signal; Phosphorylation.
 FT SIGNAL 1 19
 FT CHAIN 20 806 INTEGRIN BETA-7,
 FT DOMAIN 724 724 EXTRACELLULAR (POTENTIAL),
 FT TRANSMEM 725 745 POTENTIAL,
 FT DOMAIN 746 806 CYTOPLASMIC (POTENTIAL),
 FT DOMAIN 150 389 WFPA-LIKE,
 FT DOMAIN 478 640 4 CYSTEINE-RICH TANDEM REPEATS,
 FT REPEAT 478 526 1,
 FT REPEAT 527 565 11,
 FT REPEAT 566 604 11,
 FT REPEAT 605 640 11,
 FT REPEAT 605 640 IV.

FT DISULFID 45 476 BY SIMILARITY.
 FT DISULFID 51 61 BY SIMILARITY.
 FT DISULFID 54 91 BY SIMILARITY.
 FT DISULFID 64 80 BY SIMILARITY.
 FT DISULFID 216 223 BY SIMILARITY.
 FT DISULFID 271 311 BY SIMILARITY.
 FT DISULFID 412 428 BY SIMILARITY.
 FT DISULFID 448 688 BY SIMILARITY.
 FT DISULFID 474 478 BY SIMILARITY.
 FT DISULFID 488 500 BY SIMILARITY.
 FT DISULFID 497 537 BY SIMILARITY.
 FT DISULFID 502 511 BY SIMILARITY.
 FT DISULFID 513 527 BY SIMILARITY.
 FT DISULFID 543 548 BY SIMILARITY.
 FT DISULFID 545 574 BY SIMILARITY.
 FT DISULFID 550 559 BY SIMILARITY.
 FT DISULFID 561 566 BY SIMILARITY.
 FT DISULFID 580 585 BY SIMILARITY.
 FT DISULFID 582 613 BY SIMILARITY.
 FT DISULFID 587 596 BY SIMILARITY.
 FT DISULFID 598 605 BY SIMILARITY.
 FT DISULFID 619 624 BY SIMILARITY.
 FT DISULFID 621 666 BY SIMILARITY.
 FT DISULFID 626 635 BY SIMILARITY.
 FT DISULFID 638 641 BY SIMILARITY.
 FT DISULFID 645 654 BY SIMILARITY.
 FT MOD_RES 777 777 PHOSPHORYLATION (BY SIMILARITY).
 FT CARBOHYD 68 68 N-LINKED (GLCNAC...) (POT15).
 FT CARBOHYD 250 250 N-LINKED (GLCNAC...) (POT15).
 FT CARBOHYD 279 279 N-LINKED (GLCNAC...) (POT15).
 FT CARBOHYD 434 434 N-LINKED (GLCNAC...) (POT15).
 FT CARBOHYD 531 531 N-LINKED (GLCNAC...) (POT15).
 FT CARBOHYD 590 590 N-LINKED (GLCNAC...) (POT15).
 FT CARBOHYD 665 665 N-LINKED (GLCNAC...) (POT15).
 FT CARBOHYD 674 674 N-LINKED (GLCNAC...) (POT15).
 FT CONFLICT 81 81 A -> E (IN REF. 4).
 FT CONFLICT 81 81 S -> G (IN REF. 3).
 FT CONFLICT 124 124 MISSING (IN REF. 2).
 FT CONFLICT 538 538 S -> C (IN REF. 4).
 FT CONFLICT 557 557 S -> H (IN REF. 2).
 FT SEQUENCE 806 AA, 87411 MD, 87405 MD (GCT4).
 Query Match 2.28; Score 7; DB 1; Length 604
 Best Local Similarity 100.0%; Pref. No. 69;
 Matches 7; Conservative 0; Mismatches 0; Indels 0
 OY 71 AVGEISE 77
 DB 367 AVGEISE 373
 RESULT 33
 GIK1_DROME
 ID GIK1_DROME STANDARD; PRI: 991 AA.
 AC 003445;
 DT 15-JUL-1998 (Ref. 36, Created)
 DT 15-JUL-1998 (Ref. 36, Last sequence update)
 DT 15-JUN-2002 (Ref. 41, Last annotation update)
 DE Glutamate receptor 1 precursor (GluR1) (Kainate-selective
 DE glutamate receptor).
 GN GluR1.
 OS Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea;
 CC Insecta; Euryptera; Neoptera; Eudipterygota; Diptera; 66
 CC Muscomorpha; Ephydroidea; Trisoptillidae; Drosophila
 OX NCBI_TaxID:7227;
 RN 11
 RP SEQUENCE FROM N.A.
 RA MEDLINE:9306263; PubMed:1359540;
 RA Ullrich A., Schuster G.M., Leung E., Schless P., Schmitt
 RT Kainate-selective subunit expressed in the central nervous
 RL Proc. Natl. Acad. Sci. U.S.A. 89:10484-10488(1992).

CC -1- FUNCTION: L-Glutamate acts as an excitatory neurotransmitter at
 CC MANY SYNAPSES IN THE CENTRAL NERVOUS SYSTEM. THE POSTSYNAPTIC
 CC ACTIONS OF GLU ARE MEDIATED BY A VARIETY OF RECEPTORS THAT ARE
 CC NAMED ACCORDING TO THEIR SELECTIVE AGONISTS (BY SIMILARITY).
 CC -1- SUBUNIT: FORMS HOMOPOLYMERIC CHANNELS WHICH ARE ACTIVATED BY
 CC KAINATE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.
 CC -1- DEVELOPMENTAL STAGE: NO EXPRESSION IS SEEN IN EARLY EMBRYOGENESIS.
 CC WHEREAS HIGH EXPRESSION OCCURS IN LATE EMBRYOS DURING LARVAL
 CC DEVELOPMENT, EXPRESSION DECREASES TO UNDETECTABLE LEVELS IN LATE
 CC LARVAE, RESUMES AT THE EARLY PUPAL STAGE AND GRADUALLY INCREASES
 CC IN LATE PUPAE AND EARLY ADULT FLIES. HIGH LEVELS OF EXPRESSION
 CC COINCIDE WITH MAJOR STAGES OF NEUROGENESIS.
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by not-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.slb.ch/submit/>
 CC or send an email to license@isb.slb.ch).
 CC
 CC EMBL: M97192; AAC8575.1; ..
 CC HSSP: P14491; 1GR2.
 CC Flybase: FBga0004619; Glu-R1.
 CC InterPro: IPR001828; ANF_receptor.
 CC InterPro: IPR001320; Ion_glu_receptor.
 CC InterPro: IPR001311; SBP/qlu_receptor.
 CC Pfam: PF00060; 1lg_chan.1.
 CC Pfam: PF01094; ANF_receptor.1.
 CC ProDom: PD000500; Ion_glu_receptor.1.
 CC SMART: SM00079; PDBF.1.
 CC Receptor; Postsynaptic membrane; Ionic channel; Transmembrane;
 KW Signal: Glycoprotein.
 FT SIGNAL 1 27 POTENTIAL.
 FT CHAIN 28 991 GLUTAMATE RECEPTOR 1
 FT DOMAIN 28 611 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 612 632 POTENTIAL.
 FT TRANSMEM 711 731 POTENTIAL.
 FT TRANSMEM 895 916 POTENTIAL.
 FT DOMAIN 28 32 POTENTIAL.
 FT DOMAIN 367 371 POLY-SER.
 FT CARBOHYD 67 67 N-LINKED (GLCNAC) (POTENTIAL)
 FT CARBOHYD 145 145 N-LINKED (GLCNAC) (POTENTIAL)
 FT CARBOHYD 208 208 N-LINKED (GLCNAC) (POTENTIAL)
 FT CARBOHYD 281 281 N-LINKED (GLCNAC) (POTENTIAL)
 FT CARBOHYD 376 376 N-LINKED (GLCNAC) (POTENTIAL)
 FT CARBOHYD 385 385 N-LINKED (GLCNAC) (POTENTIAL)
 FT CARBOHYD 425 425 N-LINKED (GLCNAC) (POTENTIAL)
 FT CARBOHYD 437 437 N-LINKED (GLCNAC) (POTENTIAL)
 FT CARBOHYD 477 477 N-LINKED (GLCNAC) (POTENTIAL)
 SQ SEQUENCE 991 AA; 111555 MW; F81D5D6D614DD0D2C CRC64;
 Query Match: 2.28, Score 7, DB:1, Length: 991;
 Best local similarity: 100.0%; Pred. No. 82;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 174 Q01YQEL 180
 1111111
 185 Q01YQEL 191
 RESULT 14
 ID POLG1BHVH STANDARD: PPT: 414 AA.
 AC Q01299;
 DI 01-OCT-1996 (Ref. 34, Cytosol)
 DI 01-OCT-1996 (Ref. 34, Last sequence update)
 DI 15-JUN-2002 (Ref. 41, Last annotation update)
 DE Oocyte F-1 protein (Cytosol). Capsid protein c (Oocyte protein). Matrix
 DE protein (Envelope protein M); Major envelope protein F; Nucleocapsid

DE Proteins NS1, NS2A, NS2B, NS4A and NS4B, Protease/helicase
 DE (E:3.4.21.98) (NS3); RNA-directed RNA polymerase (E:2.7.7.48)
 DE (NS5).
 DE Tick-borne encephalitis virus (strain Hypr) (TBEV).
 CC Viruses, ssRNA positive-strand viruses, to DNA stage, Flaviviridae;
 CC Flavivirus.
 CC NCBI_TaxId:70733;
 CC
 CC [1]
 CC SEQUENCE FROM N.A.
 CC PA Walther G., Mandl C.W., Pöcker M., Holzmann H., Stiasny K.,
 CC RA Kunz C., Heinz F.X.;
 CC RT Submitted (Oct-1995) to the EMBL/Genbank/DBJ databases.
 CC RN [2]
 CC PP SEQUENCE OF 3457-3414 FROM N.A.
 CC RX MEDLINE:9130656; PubMed:1712858;
 CC RA Mandl C.W., Kunz C., Heinz F.X.;
 CC RT "Presence of poly(A) in a flavivirus: significant differences between
 CC the 3' noncoding regions of the genomic RNAs of tick-borne
 CC encephalitis virus strains";
 CC RL J. Virol. 65:4070-4077(1991).
 CC -1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
 CC HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
 CC NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
 CC precursor polyprotein, commonly with Asp or Glu in the P6
 CC position, Cys or Thr in P1 and Ser or Ala in P1'.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
 CC (RNA)(N).
 CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
 CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
 CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
 CC PROTEIN C AND NPNA.
 CC -1- MISCELLANEOUS: THE NONSTRUCTURAL PROTEINS NS1 PRESENTS TWO
 CC ALTERNATIVE CLEAVAGE SITES FOR ITS C-TERMINUS, WHICH MAY DEFINE A
 CC SOLUBLE OR A MEMBRANE-BOUND FORM OF NS1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by not-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.slb.ch/submit/>
 CC or send an email to license@isb.slb.ch).
 CC
 CC EMBL: U34292; AAB53095.1; ..
 CC EMBL: M76660; AAA47904.1; ..
 CC HSSP: P14336; 1SVB.
 CC InterPro: IPR001410; DEAD.
 CC InterPro: IPR000669; FLAVI_M.
 CC InterPro: IPR001157; FLAVI_NS1.
 CC InterPro: IPR020752; FLAVI_NS2A.
 CC InterPro: IPR000487; FLAVI_NS2B.
 CC InterPro: IPR000404; FLAVI_NS4A.
 CC InterPro: IPR001528; FLAVI_NS4B.
 CC InterPro: IPR000208; FLAVI_NS5.
 CC InterPro: IPR021122; FLAVI_capsid.
 CC InterPro: IPR050346; FLAVI_glycoprot.
 CC InterPro: IPR001850; FLAVI_helicase.
 CC InterPro: IPR002545; FLAVI_protop.
 CC InterPro: IPR002877; FtsI.
 CC InterPro: IPR001650; Helicase_C.
 CC InterPro: IPR002711; Helicase_G.
 CC Pfam: PF00869; Flavi_glycoprot.1.
 CC Pfam: PF00948; Flavi_NS1.1.
 CC Pfam: PF00949; Flavi_helicase.1.
 CC Pfam: PF00972; Flavi_NS5.1.
 CC Pfam: PF01982; Flavi_NS2B.1.
 CC Pfam: PF01003; Flavi_capsid.1.
 CC Pfam: PF01364; Flavi_M.1.
 CC Pfam: PF01905; Flavi_NS2A.1.
 CC Pfam: PF01349; Flavi_NS4B.1.
 CC Pfam: PF01450; Flavi_NS4A.1.
 CC Pfam: PF01570; Flavi_protop.1.

```

DR pfam: PF01282; FtsJ; 1.
DR pfam: PF02832; Flavyl-lysozyme C; 1.
DR pfam: PF001496; Flavi-NS1; 1.
DR pfam: PF001556; Flavi-lysozyme; 1.
DR SMART: SM00490; Helicase; 1.
KW Polyprotein; Glycoprotein; Transferrin; RNA-directed RNA polymerase;
KW Core protein; Coat protein; Envelope protein; Hydrolyase; Helicase;
KW ATP-binding; Transmembrane; Nonstructural protein;
FT INIT_MET 1
FT CHAIN 1 112
FT PROPEP 113 205
FT CHAIN 206 280
FT CHAIN 281 776
FT CHAIN 777 1128
FT CHAIN 1129 1358
FT CHAIN 1359 1489
FT CHAIN 1490 2110
FT CHAIN 2111 2259
FT CHAIN 2260 2511
FT CHAIN 2512 3414
FT NP_BIND 1688 1695
FT SITE 1779 1782
FT TRANSMEM 101 112
FT TRANSMEM 247 259
FT TRANSMEM 266 280
FT TRANSMEM 748 751
FT DISULFID 283 310
FT DISULFID 340 396
FT DISULFID 354 385
FT DISULFID 372 401
FT DISULFID 466 570
FT DISULFID 587 618
FT CARBOHYD 144 144
FT CARBOHYD 434 434
FT CARBOHYD 861 861
FT CARBOHYD 983 983
FT CARBOHYD 999 999
FT CARBOHYD 1649 1649
FT CARBOHYD 1988 1988
FT CARBOHYD 2044 2044
FT CARBOHYD 2447 2447
FT CARBOHYD 2529 2529
FT CARBOHYD 2726 2726
SQ SEQUENCE 3414 AA; 378539 MW; EC0B1A5325A08C19 CPG64.

Query Match
Best Local Similarity 100.0%; Score 7; DB 1; Length 3414.
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LIACVLY 38
DB 761 LIACVLY 767

RESULT 15
BCE1_HUMAN STANDARD; PRT; 84 AA.
AC 060756;
DI 30-MAY-2000 (Rel. 39, Created)
DI 30-MAY-2000 (Rel. 39, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE BCE-1 Protein.
GN BCE1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang Z., Huang G.S.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC

```

```

CC This SWISS-PROT entry is copyright. It is produced through
CC between the Swiss Institute of Bioinformatics and the
CC the European Bioinformatics Institute. There are no restrictions
CC use by non-profit institutions as long as its content
CC modified and this statement is not removed. Usage by
CC entities requires a license agreement (see http://www.ebi.ac.uk/seqdb/doc/seqdb.html).
CC or send an email to license@ebi.ac.uk.
DR EMBL: AF068197; AAC18942.1;
SQ SEQUENCE 84 AA; 9784 MW; D5F7B874EE12422 CPG64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 84.
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 LSFML 33
DB 71 LSFML 76

RESULT 36
ID RK21_PORPU STANDARD; PRT; 104 AA.
AC P51209;
DI 01-OCT-1996 (Rel. 34, Created)
DI 01-OCT-1996 (Rel. 34, Last sequence update)
DI 01-OCT-1996 (Rel. 34, Last annotation update)
DE Chloroplast 50S ribosomal protein L21.
GN PLE21.
OS Porphyra purpurea.
OC Chloroplast.
OC Eukaryota; Rhodophyta; Bandiophyceae; Bandiophyceae; Bandiophyceae;
OX NCBI_TaxID=2787;
RN [1]
RP SEQUENCE FROM N.A.
RA STEIN, A.;
RA Keith M.E.;
RT "Complete nucleotide sequence of the porphyra purpurea genome."
RL Plant Mol. Biol. Rep. 13:333-336(1995).
SQ SEQUENCE 104 AA; 11877 MW; 74B810785F0C2455 CPG64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 104.
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 QELRL 137
DB 92 QELRL 97

RESULT 37
ID C59A_MOUSE STANDARD; PRT; 123 AA.
AC C59A_MOUSE

```

055186;
 16-OCT-2001 (Rel. 40, last sequence update)
 16-OCT-2001 (Rel. 40, last annotation update)
 CD59A glycoprotein precursor (Membrane attack complex inhibition factor) (MACIF) (MAC-inhibitory protein) (MAC-IP) (Protectin).
 CD59A OR CD59
 Mus musculus (Mouse)
 Pukayila, Melissa; Chudalla; Crunilda; Vertebra, Eulicestomi, Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus musculus; NCBI:taxid=10090;
 11
 SEQUENCE FROM N.A.
 STRAIN=C57BL/6; TISSUE=Kidney;
 MEDLINE=97180887; PubMed=9029105;
 Powell M B, Marchbank K T, Pushmore N K, van den Berg C W, Morgan B P;
 "Molecular cloning, chromosomal localization, expression, and functional characterization of the mouse analogue of human CD59." J. Immunol. 158:1692-1702(1997).
 12
 SEQUENCE FROM N.A.
 STRAIN=129/SV;
 MEDLINE=20422499; PubMed=10965140;
 Holt D S, Powell M B, Pushmore N K, Morgan B P;
 "Genomic structure and chromosomal location of the gene encoding mouse CD59." J. Cytochrome Cell Genet. 85:264-277(2000).
 13
 SEQUENCE FROM N.A.
 STRAIN=C57BL/6; TISSUE=Medulla oblongata, kidney, and placenta;
 MEDLINE=21085660; PubMed=11217851;
 Kawai J, Shingawa A, Shibata K, Yoshino M, Itoh M, Ishi Y, Arakawa T, Hara A, Fukunishi Y, Konno H, Adachi J, Fukuda S, Aizawa K, Iizawa M, Nishi K, Kiyosawa H, Kondo S, Yamanka I, Saito T, Okazaki Y, Gotohori T, Iwano H, Kasukawa T, Saito F, Kadota K, Matsuda H A, Ashburner M, Batalov S, Casavant T, Fleischmann W, Gaasterland T, Gissi C, Kling H, Kochiya H, Kiehl P, Lewis S, Matsuo Y, Nikaike I, Pasola G, Quackenbush T, Schraml L M, Staehli F, Suzuki P, Tomita M, Wagner L, Washio T, Sakai K, Okido T, Furuno M, Aono H, Balderelli P, Birch G, Blake J, Hoffell D, Holmura N, Carninci P, de Hernaldo M P, Brownstein M J, Hult C, Fletcher C, Fujita M, Gariboldi M, Gustavich S, Hill D, Hermann M, Hume D A, Kamiya M, Lee N H, Lyons P, Marchionni L, Mashima J, Mazzarelli J, Mombereis P, Mordona P, Ping P, Ringwald M, Rodriguez I, Sakamoto N, Sasaki H, Sato K, Schoenbach C, Seya T, Shibata Y, Slouch K F, Suzuki H, Toyooka K, Wang K H, Welty C, Whitaker C, Wilming L, Wynshaw-Boris A, Yoshida K, Hasegawa Y, Kawai H, Kohlsaki S, Hayashizaki Y;
 "Functional annotation of a full-length mouse cDNA collection." Nature 409:685-690(2001).
 14
 FUNCTION: POTENT INHIBITOR OF THE COMPLEMENT MEMBRANE ATTACK COMPLEX (MAC) ACTION. ACTS BY BINDING TO THE C5 AND/OR C6 COMPLEMENTS OF THE ASSEMBLING MAC, THEREBY PREVENTING INCORPORATION OF THE MULTIPLE COPIES OF C9 REQUIRED FOR COMPLETE FORMATION OF THE OSMOTIC PORE (BY SIMILARITY).
 15
 SUBCELLULAR LOCATION: Attached to the membrane by a GPI anchor.
 16
 TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXAMINED (LIVER, KIDNEY, SPLEEN, THYMUS, BRAIN AND HEART). LOW LEVELS IN THYMUS.
 17
 ALSO EXPRESSED IN MONONUCLEAR CELLS, ERYTHROCYTES AND PLATELETS.
 18
 BARELY DETECTED IN NEUTROPHILS.
 19
 SIMILARITY: CONTAINS 1 UPART/6 DOMAIN
 20
 This SWISS-PROT entry is copyright © 1994-2002 through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.slb.ch/announcements/> or send an email to license@isb.slb.ch).
 EMBL: U60473; AAC00055.1;

EMBL: AF247652; AAC15414.1;
 EMBL: AK018146; BAB31088.1;
 EMBL: AF062748; BAB23521.1;
 EMBL: AK005507; BAB24087.1;
 DR HSSP: P1987; IERQ.
 DR MGDB: MG1.109177; CD59a.
 DR InterPro: IPR001526; IY6_0PAR.
 DR InterPro: IPR003632; IY6_0P59.
 DR Ifam: IF00221; UPAR_LY6.1.
 DR Proteome: Proteo129; LY6_0P59; 1
 DR SMART: SM00184; LIL; 1
 DR PROSITE: PS00983; LY6_0PAR; FALSE_NEG.
 KW Antigen; glycoprotein; GPI-anchor; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 96 CD59A GLYCOPROTEIN.
 FT PROPEP 97 124 PROPEP IN MATURE FORM (BY SIMILARITY).
 FT DOMAIN 24 96 UPAR/LY6.
 FT DISULFID 25 52 BY SIMILARITY.
 FT DISULFID 29 37 BY SIMILARITY.
 FT DISULFID 44 63 BY SIMILARITY.
 FT DISULFID 69 87 BY SIMILARITY.
 FT DISULFID 88 93 BY SIMILARITY.
 FT CARBOHYD 40 40 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 94 94 N-LINKED (MANN...) (POTENTIAL).
 FT LIPID 95 96 GPI-ANCHOR (BY SIMILARITY).
 SQ SEQUENCE 124 AA; 1644 MW; 1644 MW; AADVDPNP2AT374 CR64;
 Query Match 1.89, Score 6, DB 1, Length 123;
 Best Local Similarity 100.0%, Freq. No. 1.2e+02;
 Matches 61 Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 36 VLVAL. 41
 DB 110 VLVAL. 115
 ID P271_SVNY3 STANDARD; PPT: 124 AA
 RL21_SVNY3
 AC P74266;
 DT 21-NOV-1997 (Rel. 35, last sequence update)
 DT 01-NOV-1997 (Rel. 35, last annotation update)
 DT 16-OCT-2001 (Rel. 40, last sequence update)
 DT 50S ribosomal protein L21.
 DE PPT of PPT21 GP SLP1578.
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chlorococcales; Synechocystis.
 OX NCBI:taxid=1148;
 RN 111
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97051201; PubMed=8905231;
 RA Kunko T, Sato S, Kolum H, Tanaka A, Asanuma E, Nakamura Y, Miyajima N, Hirosewa M, Sugita M, Sasamoto S, Kimura T, Okumori T, Matsuno A, Morita A, Nakazaki N, Naruo K, Okamura S, Shimo S, Takeuchi C, Wada T, Matsumoto A, Yamada M, Yasuda M, Tabata S;
 "Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. II. Sequence determination of the full-length genes and assignment of potential protein-coding regions." J. Mol. Biol. 310:9-146(1996).
 12
 FUNCTION: THIS PROTEIN BINDS TO 23S RIBOSOMAL RNA IN THE PRESENCE OF PROTEIN L20 (BY SIMILARITY).
 13
 OF PROTEIN L20 (BY SIMILARITY).
 14
 SIMILARITY: BELONGS TO THE L21 FAMILY OF RIBOSOMAL PROTEINS.
 15
 This SWISS-PROT entry is copyright © 1994-2002 through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.slb.ch/announcements/> or send an email to license@isb.slb.ch).
 EMBL: D90913; BAB18360.1;

```

DR InterPro: IPR001787; Ribosomal_L21P.
DR Pfam: PF00829; Ribosomal_L21P; 1.
DR ProDom: PD003604; Ribosomal_L21P; 1.
DR TrEMBL: TIGR00061; L21; 1.
DR PROSITE: PS01169; RIBOSOMAL_L21; 1.
DR Ribosomal protein; rRNA-binding; Complete proteome.
SQ SEQUENCE 124 AA; 13669 MW; 0201A7C57AFEB0D CRC64;

Query Match 1.88; Score 6; DB 1; Length 124;
Best Local Similarity 100.0%; Pred. No. 1.3e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 QELIRL 114
DB 92 QELIRL 97

RESULT 49
AAA_TRIAL STANDARD; PRT; 125 AA.
AC P81133.
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alboagregin A subunit 3.
OS Trimeresurus albolabris (White-lipped pit viper),
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Leiodosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
Viperidae; Crotalinae; Trimeresurus.
OX NCBI_TaxID:8765;
RN [1]
RP TISSUE:Venom;
RX MEDLINE=98189535; PubMed=9531050;
RA Kowalska M.A., Tan L., Holt J.C., Peng M., Karocwski J.,
RT Calvete J.J., Niewiarowski S.;
*Alboagregin A and B. Structure and interaction with human
RT platelets.*;
RL Thromb Haemost. 79:609-613(1998).
CC -!- FUNCTION: BINDS TO PLATELET GPIIb/IIIa RECEPTOR SYSTEM AND STIMULATES
AGGREGATION.
CC -!- SUBUNIT: HETEROETRAMER OF THE SUBUNITS 1, 2, 3 AND 4,
CC -!- DISULFIDE-LINKED.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN
HSSP: P23807.11XX.
DR InterPro: IPR001304; LECTIN_C.
DR Pfam: PF00059; LECTIN_C; 1.
DR SMART: SM00034; CLECT; 1.
DR PROSITE: PS00615; C-TYPE_LECTIN_1; FALSE_NEG.
DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
KW Venom; LECTIN.
FT DOMAIN 3 123 C-TYPE LECTIN (LONG FORM).
FT DISULFID 4 15 BY SIMILARITY.
FT DISULFID 32 121 BY SIMILARITY.
FT DISULFID 98 113 BY SIMILARITY.
SQ SEQUENCE 125 AA; 14798 MW; CAFE24C09BDF3293 CRC64;

Query Match 1.88; Score 6; DB 1; Length 125;
Best Local Similarity 100.0%; Pred. No. 1.3e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 WMGLSD 258
DB 67 WMGLSD 72

RESULT 40
C59B_MOUSE STANDARD; PRT; 129 AA.
AC P58019.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

```

```

DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE CD59b glycoprotein precursor (Membrane attack complex in-
DE factor) (MACIF) (MAC-inhibitory protein) (MAC-IP) (trite
GN CD59b).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eute-
OC leomammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mu-
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=129/Sv.
RX MEDLINE=20405635; PubMed 10946279;
RA Qian Y.-M., Qin X., Miwa T., Sun X., Halperin J.A., Song
RT *Identification and functional characterization of a nov-
RT the mouse terminal complement inhibitor CD59.*;
RL J. Immunol. 165:2528-2534(2000).
CC -!- FUNCTION: POTENT INHIBITOR OF THE COMPLEMENT MEMBRANE
COMPLEX (MAC) ACTION. ACTS BY BINDING TO THE CR AN/
CC COMPLEMENTS OF THE ASSEMBLING MAC, THEREBY PREVENTING
CC INCORPORATION OF THE MULTIPLE COPIES OF C9 REQUIRED
CC FORMATION OF THE OSMOTIC PORE (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a
CC -!- TISSUE SPECIFICITY: SPECIFICALLY IN TESTIS.
DR HSSP: P13987; ICDS.
DR MGD: MG1:1888966; Cd59b.
DR InterPro: IPR001526; LY6_UPAR.
DR InterPro: IPR003632; LY-6_CD59.
DR Pfam: PF00021; UPAR_LY6; 1.
DR ProDom: PD003128; LY-6_CD59; 1.
DR SMART: SM00134; LY; 1.
DR PROSITE: PS00683; LY6_UPAR; FALSE_NEG.
KW Antigen; Glycoprotein; GPI-anchor; Signal.
FT SIGNAL 1 23
FT CHAIN 24 107 CD59b GLYCOPROTEIN.
FT PROPEP 108 129 REMOVED IN MATURE FORM (BY
FT DOMAIN 24 107 UPAR/LY6.
FT DISULFID 26 49 BY SIMILARITY.
FT DISULFID 23 36 BY SIMILARITY.
FT DISULFID 42 62 BY SIMILARITY.
FT DISULFID 68 86 BY SIMILARITY.
FT DISULFID 87 92 BY SIMILARITY.
FT CARBOHYD 39 39 N-LINKED (GLCNAC...) (PRO-
FT LIPID 107 107 GPI-ANCHOR (BY SIMILARITY);
SQ SEQUENCE 129 AA; 14266 MW; 2561906A9A9C9CCE CRC64;

Query Match 1.88; Score 6; DB 1; Length 129;
Best Local Similarity 100.0%; Pred. No. 1.3e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 VLVAIL 41
DB 120 VLVAIL 125

RESULT 41
CYCP_RHOGE STANDARD; PRT; 129 AA.
AC P00142.
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Cytochrome c.
OS Rhodocyclus gelatinosus (Rhodospirillum rubrum).
OC Bacteria; Proteobacteria; beta subdivision; Comamonada-
OC rubrivivax.
OX NCBI_TaxID=28068;
RN [1]
RP SEQUENCE.
RX MEDLINE=79199606; PubMed=221823;
RA Ambler R.P., Meyer T.E., Kamen M.D.;
RT *Anomalies in amino acid sequences of small cytochromes
RT cytochromes c' from two species of Purple Phototrophic

```



```

CC Elasmobranchii: Squala, Hypnosquala; Pisticidae: Batoidae;
CC Torpediniformes: Torpedinidae; Torpedinidae: Torpedo.
CX NCBI_TaxID=7788;
RN [1]
RP SEQUENCE:
RA MEDLINE=90074179; PubMed=2500465;
RA Huber F., Braunltzer G.;
RT "The primary structure of electric ray hemoglobin (Torpedo marmorata).
RT Bohr effect and phosphate interaction.";
RL Biol. Chem. Hoppe-Seyler 370:841-848(1989)
CC -1- FUNCTION: Involved in oxygen transport from the lung to the
CC various peripheral tissues.
CC -1- SUBUNIT: Heterotetramer of two alpha chains and two beta chains.
CC -1- TISSUE SPECIFICITY: Red blood cells.
CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.
DR PIR: S05418; HARYM.
DR HSSP: p56691; ICG5.
DR InterPro: IPR000971; Globin.
DR Pfam: PF00042; globin; 1.
DR PROSITE: PS01033; GLOBIN; 1.
KW Heme; oxygen transport; Transport; Erythrocyte.
FT METAL. 59 59 IRON (HEME DISTAL LIGAND);
FT METAL. 88 88 IRON (HEME PROXIMAL LIGAND);
SQ SEQUENCE 141 AA; 15866 MW; 31BA7A1756FC06B7 CPO64;

Query Match 1.88; Score 6; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 HGALVL 25
    |||||
DB 59 HGALVL 64

RESULT 45
HBF1_URECA STANDARD; PRT; 141 AA.
ID HBF1_URECA
AC P06148;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hemoglobin F-1.
OS Urechis caupo (Innkeeper worm) (Spoonworm).
OC Eukaryota, Metazoa, Echinua, Xenopneustae; Urechidae; Urechis.
CX NCBI_TaxID=6431;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=87057334; PubMed=3782128;
RA Garey J.R., Riggs A.F.;
RT "The hemoglobin of Urechis caupo. The cDNA-derived amino acid
RT sequence.";
RL J. Biol. Chem. 261:16446-16450(1986).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RA MEDLINE=92384990; PubMed=1515107;
RA Kolatkar P.R., Ernst S.R., Hackert M.L., Ogata C.M.,
RA Hendrickson W.A., Merril E.A., Phizackerley R.P.;
RT "Structure determination and refinement of homotetrameric hemoglobin
RT from Urechis caupo at 2.5-A resolution.";
RL Acta Crystallogr. B 48:191-199(1992).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RA MEDLINE=94180394; PubMed=813523;
RA Kolatkar P.R., Hackert M.L., Riggs A.F.;
RT "Structural analysis of Urechis caupo hemoglobin.";
RL J. Mol. Biol. 237:87-97(1994).
CC -1- FUNCTION: HEMOGLOBIN F-1 APPEARS TO FUNCTION IN STORAGE, FATHER
CC THAN TRANSPORT OF OXYGEN.
CC -1- SUBUNIT: Homotetramer.
CC -1- MISCELLANEOUS: HEMOGLOBIN F-1 BINDS OXYGEN NONCOOPERATIVELY AND
CC ALMOST INDEPENDENTLY OF PH.
CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.

```

```

CC This SWISS-Prot entry is copyright. It is produced through
CC between the Swiss Institute of Bioinformatics and the
CC the European Bioinformatics Institute. There are no fees
CC use by non-profit institutions as long as its content
CC modified and this statement is not removed. Usage by
CC entities requires a license agreement (see http://www.ebi.ac.uk
CC or send an email to license@sb.slb.ch).
CC -----
CC EMBL: J02624; AAA30331.1;
CC PIR: A25577; A25537.
CC PDB: 1TTH; 3I-OCY-93.
DR InterPro: IPR000971; Globin.
DR Pfam: PF00042; globin; 1.
DR PROSITE: PS01033; GLOBIN; 1.
KW Heme; oxygen transport; Transport; Erythrocyte.
KW 3D-structure.
FT INT_MET 0 0 IRON (HEME PROXIMAL LIGAND)
FT METAL 94 94
FT HELIX 4 17
FT TURN 18 18
FT HELIX 19 36
FT HELIX 38 43
FT TURN 45 49
FT HELIX 52 57
FT HELIX 59 77
FT TURN 78 82
FT HELIX 83 88
FT TURN 89 90
FT HELIX 91 95
FT TURN 96 98
FT HELIX 101 118
FT HELIX 123 138
FT TURN 139 140
SQ SEQUENCE 141 AA; 15062 MW; 2F68H507903E646A CPO64;

Query Match 1.88; Score 6; DB 1; Length 141
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 AGVLVA 39
    |||||
DB 133 AGVLVA 138

```

Search completed: December 7, 2002, 11:16:09
Job time : 28 secs

100

GenCore version 5.1.1
Copyright (c) 1993 - 2002 CompuGen Ltd

OM protein - protein search, using sw model

Run on: December 7, 2002, 11:08:20, Search time: 33 Seconds
(without alignments)
2029, 252 Million cell updates/sec

Title: US-09-831-458A-12

Perfect score: 325
Sequence: 1 MSDSKEPRVQULGLGLJH PVDVNYWICKKPAATFPE 325

Scoring table:

	OLIGO
Gapop 60 0	Gapext 60 0

Searched: 671580 seqs. 206047115 residues

Word Size :

Total number of hits satisfying chosen parameters: 33422

Minimum PB seq length: 0
Maximum PB seq length: 2000000000

Post-processing: Listing first 45 summaries

```

Database :
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriap:*
17: sp.archaeop:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query No.	Score	Match	Length	DB	ID	Description
	1	168	51.7	296	4	Q9Y2A8	Q9Y2A8 homo sapien
	2	165	51.1	376	4	Q9BXS3	Q9BXS3 homo sapien
	3	166	51.1	399	4	Q9H2X3	Q9H2X3 homo sapien
	4	153	47.1	272	4	Q96M94	Q96M94 homo sapien
	5	153	47.1	399	4	Q9H2D9	Q9H2D9 homo sapien
	6	138	42.5	263	4	Q96DQ3	Q96DQ3 homo sapien
	7	138	42.5	332	4	Q96DQ5	Q96DQ5 homo sapien
	8	125	38.5	251	4	Q96QY4	Q96QY4 homo sapien
	9	120	36.9	232	4	Q96QY6	Q96QY6 homo sapien
	10	97	29.8	216	4	Q9H8F0	Q9H8F0 homo sapien
	11	78	24.0	148	4	Q96DQ7	Q96DQ7 homo sapien
	12	78	24.0	297	4	Q96QJ8	Q96QJ8 homo sapien
	13	78	24.0	312	4	Q96DQ7	Q96DQ7 homo sapien
	14	78	24.0	334	4	Q96DQ9	Q96DQ9 homo sapien
	15	78	24.0	360	4	Q96DQ4	Q96DQ4 homo sapien
	16	78	24.0	380	4	Q96DQ5	Q96DQ5 homo sapien

[illegible]

A1.13MFIN15

Query Match	51.7%	Score 166	19.4	Length 276
Best Local Similarity	100.0%	Prod. No. 1	9e-165	
Matches	100%	Conservative	0	Mismatches
			0	Indels
Q9Y2A8	PRELIMINARY		276 AA	
AC Q9Y2A8				
DT 01-NOV-1999	(TRMBLrel. 12, Created)			
DT 01-NOV-1999	(TRMBLrel. 12, last annotation update)			
DT 01-JUN-2002	(TRMBLrel. 21, last annotation update)			
DE	Type II membrane protein similar to HIV gp120-binding site lectin.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo			
OX	NCBI_Taxid:9606;			
FN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=LIVER;			
FX	MP0LINE=94173880; Pubmed=10072769;			
RA	Koyoyama-Kobayashi M., Yamaguchi I., Sasaki S., Kato S.,			
RT	"cell surface form of cDNAs encoding putative type II membrane protein			
RL	gene 228:151-167(1999)."			
DR	EMBL; AB015629; BAA76496.1;			
DR	InterPro: IP001304; lectin_C,			
DR	Pfam: PF00059; lectin_C1.1;			
SV	SEQUENCE 276 AA; 33527 MW; 74228484EF20B4F67C04;			

```

Db 123 LPEKSLGCIYQELTFLKAAVCELPKPSKICPIYQELTFLKAAVCELP 168
|||||
RESULT 2
Q9BXK3 PRELIMINARY: PRT: 376 AA
AC Q9BXK3:
DI 01-JUN-2001 (TREMblrel: 17, created)
DI 01-JUN-2001 (TREMblrel: 17, last sequence update)
DI 01-JUN-2002 (TREMblrel: 21, last annotation update)
DE L-SIGN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_LexID-9606;
RN 111
RP SEQUENCE FROM N.A.
RX MEDLINE:21157496; PubMed 11257134;
RA Bashirova A.A., Gelfandov T.R.H., van Duinheven G.C.F.,
RA van Vliet S.J., Ellering J.B.G., Martin M.P., Wu L., Martin T.H.,
RA Viebig N., Knolle P.A., Kewajamani V.N., van Kooyk Y., Carrington M.;
KT "A dendritic cell-specific intercellular adhesion molecule 3-grabbing
KT nonintegrin (DC-SIGN)-related protein is highly expressed on human
KT liver sinusoidal endothelial cells and promotes HIV-1 infection.";
RL J. Exp. Med. 193:671-678(2001)
DR EMBL: AF290887; AAK20998.1; -.
DR HSP: P22897; IEG3.
DR InterPro: IPR002353; AntiFcecell.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C_1.
DR PRINTS: PR00366; ANTIPEFEZFI1.
DR SMART: SM00034; CLECT_1.
DR PROSITE: PS00615; G_TYPE_LECTIN_1; DNKNOWN_1.
DR PROSITE: PS50641; G_TYPE_LECTIN_2; 1.
SQ SEQUENCE 376 AA: 42724 MW: 36319991F23BE90 CQC64;

Query Match 51.1%; Score 166; DB 4; Length 376;
Best Local Similarity 100.0%; Pred. No. 2,78-153;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 LKAAVCELPKPSKICPIYQELTFLKAAVCELPKPSKICPIYQELTFLKAAVCELP 219
Db 211 LKAAVCELPKPSKICPIYQELTFLKAAVCELPKPSKICPIYQELTFLKAAVCELP 270
QY 220 TACQEVRAQIVVVKTAEPGNPIQIQTSPNSNPFSSWAGLSDLNCEJTWQVWDSPLSPSPQR 279
Db 271 TACQEVRAQIVVVKTAEPGNPIQIQTSPNSNPFSSWAGLSDLNCEJTWQVWDSPLSPSPQR 330
QY 280 YNNSGEFNNSGNEDCAEFSGSGMNDKLDVNTWICKKPAACFDE 325
Db 331 YNNSGEFNNSGNEDCAEFSGSGMNDKLDVNTWICKKPAACFDE 376

RESULT 3
Q9H2X3 PRELIMINARY: PRT: 399 AA.
AC Q9H2X3:
DI 01-MAR-2001 (TREMblrel: 16, created)
DI 01-MAR-2001 (TREMblrel: 16, last sequence update)
DI 01-JUN-2002 (TREMblrel: 21, last annotation update)
DE Probable mannose binding C-type lectin DC-SIGN2 type I
DE isoform.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_LexID-9606;
RN 111
RP SEQUENCE FROM N.A.
RX MEDLINE:23432267; PubMed 10757799;
RA Solliou E.J., Barten R., Trowsdale J.;
DR EMBL: AV042245; AAK91860.1; -.
SQ SEQUENCE 376 AA: 42724 MW: 36319991F23BE90 CQC64;

```

```

RT 19913.7;
P: 1 Immunol. 165:2927-2942(2000)
RN 121
RP SEQUENCE FROM N.A.
RA Solliou E.J.;
RN Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RX PubMed 11377487;
RA Munnadi S., Calado G., Lam L., Heetle A., Telles V., Begum K.,
RA Jimenez F., Ahuja S.S., Ahuja S.K.;
RT "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN1) and DC-SIGN2
RT Isoforms: Inter-individual Variation in Expression of DC-SIGN
RT Transcripts.";
RL J. Biol. Chem. 275:33396-33212(2000).
DR EMBL: AF209481; AAC13815.2; -.
DR EMBL: AF209480; AAC13815.2; J01NEO.
DR EMBL: AV042244; AAK91859.1; -.
DR HSP: P22897; IEG3.
DR InterPro: IPR002353; AntiFcecell.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C_1.
DR PRINTS: PR00366; ANTIPEFEZFI1.
DR SMART: SM00034; CLECT_1.
DR PROSITE: PS00615; G_TYPE_LECTIN_1; DNKNOWN_1.
DR PROSITE: PS50641; G_TYPE_LECTIN_2; 1.
DR EMBL: AF290887; AAK20998.1; -.
SQ SEQUENCE 399 AA: 45350 MW: 06A6709372AFA593 CQC64;

Query Match 51.1%; Score 166; DB 4; Length 399;
Best Local Similarity 100.0%; Pred. No. 2,78-153;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 LKAAVCELPKPSKICPIYQELTFLKAAVCELPKPSKICPIYQELTFLKAAVCELP 219
Db 234 LKAAVCELPKPSKICPIYQELTFLKAAVCELPKPSKICPIYQELTFLKAAVCELP 293
QY 220 TACQEVRAQIVVVKTAEPGNPIQIQTSPNSNPFSSWAGLSDLNCEJTWQVWDSPLSPSPQR 279
Db 294 TACQEVRAQIVVVKTAEPGNPIQIQTSPNSNPFSSWAGLSDLNCEJTWQVWDSPLSPSPQR 353
QY 280 YNNSGEFNNSGNEDCAEFSGSGMNDKLDVNTWICKKPAACFDE 325
Db 354 YNNSGEFNNSGNEDCAEFSGSGMNDKLDVNTWICKKPAACFDE 399

RESULT 4
Q969M4 PRELIMINARY: PRT: 272 AA.
AC Q969M4:
DI 01-DEC-2001 (TREMblrel: 19, created)
DI 01-DEC-2001 (TREMblrel: 19, last sequence update)
DI 01-DEC-2001 (TREMblrel: 19, last annotation update)
DE MDC-SIGN2 type IV isoform (MDC-SIGN2 type IVT isoform).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_LexID-9606;
RN 111
RP SEQUENCE FROM N.A.
RX PubMed 11377487;
RA Munnadi S., Calado G., Lam L., Heetle A., Telles V., Begum K.,
RA Jimenez F., Ahuja S.S., Ahuja S.K.;
RT "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN1) and DC-SIGN2
RT Isoforms: Inter-individual Variation in Expression of DC-SIGN
RT Transcripts.";
RL J. Biol. Chem. 275:33396-33212(2000).
DR EMBL: AV042245; AAK91860.1; -.
DR EMBL: AV042244; AAK91859.1; -.
SQ SEQUENCE 272 AA: 30481 MW: 33935225E541F5E CQC64;

```

```

Query Match          47 18; Score 153; DB 4; Length 272;
Best Local Similarity 100.0%; Pred. No. 66-150;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 GCLGHALVLDLSPMLAGVLAIVLVQSKVPSSLSQESQEDATYQNTQTKAAVCEL 75
      |||||||
Db 44 GCLGHALVLDLSPMLAGVLAIVLVQSKVPSSLSQESQEDATYQNTQTKAAVCEL 103
      |||||||

QY 76 SEKSKLOEIVQELTQTKAAVGEIPKSKLOEIVQELTRKAAVGEIPKSKLOEIVQELT 135
      |||||||
Db 104 SEKSKLOEIVQELTQTKAAVGEIPKSKLOEIVQELTRKAAVGEIPKSKLOEIVQELT 163
      |||||||

QY 136 RLKAAVGEIPKSKLOEIVQELTRKAAVGEIP 168
      |||||||
Db 164 RLKAAVGEIPKSKLOEIVQELTRKAAVGEIP 196
      |||||||

RESULT 5
Q9H209 PRELIMINARY; PRT; 399 AA.
AC 09H209;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Probable mannose binding C-type lectin DC-SIGNR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;
OC NCBI_Taxid=9606;
LN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=20432267; PubMed=10975799;
RA Soilleux E.J.; Batten R.; Trowsdale J.;
RT "DC-SIGN: a related gene, DC-SIGNR, and CD23 form a cluster on
      1p13."
RL J. Immunol. 165:2937-2942(2000).
RN 12
RP SEQUENCE FROM N.A.
RA Soilleux E.J.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DDBI databases.
DR EMBL: AF245219; AAG13848.2;
DR HSSP: P22897; IEGG.
DR InterPro: IPR002353; Antifreeze1.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C.1
DR PRINTS: PR00356; ANTIFREEZE11.
DR SMART: SM00034; CLECT_1.
DR PROSITE: PS50041; C_TYPE_LECTIN_2; 1.
DR KW
SU SEQUENCE 399 AA; 45357 MW; FRC999F72AE6589 CPO64;

Query Match          47 18; Score 153; DB 4; Length 399;
Best Local Similarity 100.0%; Pred. No. 8-56-150;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 GCLGHALVLDLSPMLAGVLAIVLVQSKVPSSLSQESQEDATYQNTQTKAAVCEL 75
      |||||||
Db 44 GCLGHALVLDLSPMLAGVLAIVLVQSKVPSSLSQESQEDATYQNTQTKAAVCEL 103
      |||||||

QY 76 SEKSKLOEIVQELTQTKAAVGEIPKSKLOEIVQELTRKAAVGEIPKSKLOEIVQELT 135
      |||||||
Db 104 SEKSKLOEIVQELTQTKAAVGEIPKSKLOEIVQELTRKAAVGEIPKSKLOEIVQELT 163
      |||||||

QY 136 RLKAAVGEIPKSKLOEIVQELTRKAAVGEIP 168
      |||||||
Db 164 RLKAAVGEIPKSKLOEIVQELTRKAAVGEIP 196
      |||||||

RESULT 6
Q960P3 PRELIMINARY; PRT; 263 AA.
AC 0960P3;

```

```

DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE SDC SIGN2 type III isoform.
GN CD209L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;
OC NCBI_Taxid=9606;
LN 11
RP SEQUENCE FROM N.A.
RX PubMed=11337487;
RA Mummidi S.; Catano G.; Lam T.; Hoelle A.; Telles V.; Boj
RA Jimenez F.; Ahuja S.S.; Ahuja S.K.;
RT "Extensive repertoire of Membrane-bound and Soluble Dendritic
RT specific ICAM-3-grabbing Neutrophin 1 (DC-SIGN) and D-
RT Isoforms. Inter-individual Variation in Expression of D-
RT Transcripts."
RL J. Biol. Chem. 276:33196-33212(2001).
RL EMBL: AY042440; AAK91865.1;
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C.1
DR PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE: PS50041; C_TYPE_LECTIN_2; 1.
SU SEQUENCE 263 AA; 30102 MW; 761176295424CF CPO64;

Query Match          42 58; Score 138; DB 4; Length 272;
Best Local Similarity 100.0%; Pred. No. 26-134;
Matches 138; Conservative 0; Mismatches 0; Indels 0;

QY 188 FPLCPHPKPKMTFFQNGYFENSSGPKMHHSVTAQGVKALVYIKAFEL 272
      |||||||
Db 126 EKLCPHPKPKMTFFQNGYFENSSGPKMHHSVTAQGVKALVYIKAFEL 272
      |||||||

QY 248 SNKFSWMLSDLNQGTQWVWVWVSPISPSFORVWNSGPPNNSGNDCAFFS 272
      |||||||
Db 186 SNKFSWMLSDLNQGTQWVWVWVSPISPSFORVWNSGPPNNSGNDCAFFS 272
      |||||||

QY 308 DVDNTWCKKPPACGRDE 325
      |||||||
Db 246 EVNRYWCKKPPACGRDE 263
      |||||||

RESULT 7
Q960P5 PRELIMINARY; PRT; 332 AA.
AC 0960P5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE SDC SIGN2 type I isoform.
GN CD209L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;
OC NCBI_Taxid=9606;
LN 11
RP SEQUENCE FROM N.A.
RX PubMed=11337487;
RA Mummidi S.; Catano G.; Lam T.; Hoelle A.; Telles V.; Boj
RA Jimenez F.; Ahuja S.S.; Ahuja S.K.;
RT "Extensive repertoire of Membrane-bound and Soluble Dendritic
RT specific ICAM-3-grabbing Neutrophin 1 (DC-SIGN) and D-
RT Isoforms. Inter-individual Variation in Expression of D-
RT Transcripts."
RL J. Biol. Chem. 276:33196-33212(2001).
RL EMBL: AY042438; AAK91863.1;
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C.1
DR PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE: PS50041; C_TYPE_LECTIN_2; 1.
SU SEQUENCE 332 AA; 37954 MW; 761176295424CF CPO64;

```

Query Match 42.5%; Score 138; DB 4; Length 332;
 Best Local Similarity 100.0%; Pred. No. 2.5e-134;
 Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 188 ERLPHPTKFWI+PQSN+YFMSNSQ+NMH+SVIA+QEVPAALVVIKIAEKNPILQIQLSP 247
 |||||
 DB 195 ERLPHPTKFWI+PQSN+YFMSNSQ+NMH+SVIA+QEVPAALVVIKIAEKNPILQIQLSP 254
 |||||

UY 248 SRFESWME+SDLNPGSTGOWWPGSP+SPSPQYKNGSGPPNNGNHYVAEFGSGMNINFG 307
 |||||
 DB 255 SRFESWME+SDLNPGSTGOWWPGSP+SPSPQYKNGSGPPNNGNHYVAEFGSGMNINFG 314
 |||||

UY 308 DVDNWKCKKPAACPFDE 325
 |||||
 DB 315 DVDNWKCKKPAACPFDE 332
 |||||

RESULT 8
 Q960P4 PRELIMINARY; PRT; 251 AA.

AC Q960P4
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE SDC-SIGN2 type II isoform.
 GN CD209L.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mummidi S., Catalan G., Lam L., Hoeftle A., Telles V., Begum K.,
 RA Jimenez F., Ahuja S.S., Ahuja S.K.;
 RT "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
 RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
 RT Isoforms, Inter-Individual Variation in Expression of DC-SIGN
 RT Transcripts."
 RL J. Biol. Chem. 276:33196-33212(2001)
 DR EMBL: AY042239; AAK91864.1;
 SO SEQUENCE 251 AA; 28373 MW; 86F69F6C6D055DCE CRC64.

Query Match 38.5%; Score 125; DB 4; Length 251;
 Best Local Similarity 100.0%; Pred. No. 5.8e-171;
 Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 44 VSKVSSLSQGEQDAIYQNTOLKAAGVPISEKSLQETIYQELTOLKAAGVPEPSK 104
 |||||
 DB 51 VSKVSSLSQGEQDAIYQNTOLKAAGVPISEKSLQETIYQELTOLKAAGVPEPSK 110
 |||||

UY 104 LOETIYQELTRKAAGVPEPSKLOETIYQELTRKAAGVPEPSKLOETIYQELTRKAA 164
 |||||
 DB 111 LOETIYQELTRKAAGVPEPSKLOETIYQELTRKAAGVPEPSKLOETIYQELTRKAA 170
 |||||

UY 164 VSELP 168
 |||||
 DB 171 VSELP 175
 |||||

RESULT 9
 Q960P6 PRELIMINARY; PRT; 232 AA.

AC Q960P6
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE MUC-SIGN2 type VI isoform.
 GN CD209L.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Mummidi S., Catalan G., Lam L., Hoeftle A., Telles V., Begum K.,
 RA Jimenez F., Ahuja S.S., Ahuja S.K.;
 RT "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
 RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
 RT Isoforms, Inter-Individual Variation in Expression of DC-SIGN
 RT Transcripts."
 RL J. Biol. Chem. 276:33196-33212(2001).
 DR EMBL: AY042237; AAK91862.1;
 DR InterPro: IPR001304; Icdin_C1.
 DR Pfam: PF00059; Icdin_C1.
 SO SEQUENCE 232 AA; 26255 MW; 19C1B1502451B72 CRC64;

Query Match 36.9%; Score 120; DB 4; Length 232;
 Best Local Similarity 100.0%; Pred. No. 8.2e-116;
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 16 CCLGHCAVIQLISPMILACVIALVGVSKVPSLSQGEQDAIYQNTOLKAAGVCEL 75
 |||||
 DB 44 CCLGHCAVVIQLISPMILAGVVALIVGVSKVPSLSQGEQDAIYQNTOLKAAGVCEL 103
 |||||

UY 76 SEKSKLQETIYQELTOLKAAGVPEPSKLOETIYQELTRKAAVPEPSKLOETIYQELT 135
 |||||
 DB 104 SEKSKLQETIYQELTOLKAAGVPEPSKLOETIYQELTRKAAVPEPSKLOETIYQELT 163
 |||||

RESULT 10
 Q9H8F0 PRELIMINARY; PRT; 216 AA.

AC Q9H8F0
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
 DE CDNA FLV1006118, clone pLACF200002, highly similar to Homo sapiens
 DE type II membrane protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE-PLACENTA;
 RA Isaksson, A., Hagström, L., Saitama, T., Otsuki, T., Sasaki, T.,
 RA Nishikawa, T., Nagai, K., Srinane, S., Takahashi-Fujii, A., Hara, H.,
 RA Tanabe, T., Nomura, Y., Tohyama, S., Komai, F., Hara, R., Takeuchi, K.,
 RA Arita, M., Nakamura, T., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J.,
 RA Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuno, Y., Oshima, A.;
 RT "NEBO human cDNA sequencing project."
 RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF024750; BAB14667.1;
 SO SEQUENCE 216 AA; 24438 MW; 4A550078BE2C3CB8 CRC64;

Query Match 29.8%; Score 97; DB 4; Length 216;
 Best Local Similarity 100.0%; Pred. No. 5.2e-92;
 Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 72 VGEISKSLQETIYQELTOLKAAGVPEPSKLOETIYQELTRKAAVPEPSKLOETIY 141
 |||||
 DB 44 VGEISKSLQETIYQELTOLKAAGVPEPSKLOETIYQELTRKAAVPEPSKLOETIY 104
 |||||

UY 133 VGEISKSLQETIYQELTOLKAAGVPEPSKLOETIYQELTRKAAVPEPSKLOETIY 168
 |||||
 DB 104 VGEISKSLQETIYQELTOLKAAGVPEPSKLOETIYQELTRKAAVPEPSKLOETIY 140
 |||||

RESULT 11
 Q960P7 PRELIMINARY; PRT; 148 AA.

AC Q960P7
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)

DT 01-DEC-2001 (TfEMBLrel. 19, last annotation update)
 DE SDC-SIGNIB type IV isoform.
 GN CD209.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=1137487;
 RA Mummidi S., Catano G., Lam L., Hoelle A., Telles V., Begun K.,
 RJ Jimenez F., Ahuja S.S., Ahuja S.K.;
 RT *Extensive repertoire of Membrane-bound and Soluble pendritic Cell-
 specific ICAm-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
 RT Isoforms. Inter-individual Variation in Expression of DC-SIGN
 RT Transcripts.*;
 RL J. Biol. Chem. 276:33196-33212(2001).
 DR EMBL: AY042333; AAK91858.1;
 SO SEQUENCE 148 AA; 16608 MW; 5FDFD84AD6DA1AF7 CRC64;

Query Match 24.0%; Score 78; DB 4; Length 148;
 Best Local Similarity 100.0%; Pred No. 1.8e-72;
 Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QDAIYONLTOLKAAGVGLSKLOEIYQELTOLKAAGVGLPEKSKLOEIYQELTIRLAKA 117
 DB 50 QDAIYONLTOLKAAGVGLSKLOEIYQELTOLKAAGVGLPEKSKLOEIYQELTIRLAKA 109
 QY 118 VGLPEKSKLOEIYQELT 135
 DB 110 VGLPEKSKLOEIYQELT 127

RESULT 12

Q96Q98 PRELIMINARY; PRT; 297 AA.
 ID Q96Q98;
 AC Q96Q98;
 DT 01-DEC-2001 (TfEMBLrel. 19, Created)
 DT 01-DEC-2001 (TfEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TfEMBLrel. 21, Last annotation update)
 DE SDC-SIGNIB type III isoform.
 GN CD209.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=1137487;
 RA Mummidi S., Catano G., Lam L., Hoelle A., Telles V., Begun K.,
 RJ Jimenez F., Ahuja S.S., Ahuja S.K.;
 RT *Extensive Repertoire of Membrane-bound and Soluble pendritic Cell-
 specific ICAm-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
 RT Isoforms. Inter-individual Variation in Expression of DC-SIGN
 RT Transcripts.*;
 RL J. Biol. Chem. 276:33196-33212(2001).
 DR EMBL: AY042332; AAK91857.1;
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF00059; Lectin_C.1;
 SO SEQUENCE 297 AA; 31873 MW; C8C0951B915FC5BF CRC64;

Query Match 24.0%; Score 78; DB 4; Length 297;
 Best Local Similarity 100.0%; Pred No. 3.4e-72;
 Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QDAIYONLTOLKAAGVGLSKLOEIYQELTOLKAAGVGLPEKSKLOEIYQELTIRLAKA 117
 DB 50 QDAIYONLTOLKAAGVGLSKLOEIYQELTOLKAAGVGLPEKSKLOEIYQELTIRLAKA 109
 QY 118 VGLPEKSKLOEIYQELT 135
 DB 110 VGLPEKSKLOEIYQELT 127

RESULT 13

Q96Q07 PRELIMINARY; PRT; 412 AA.
 ID Q96Q07;
 AC Q96Q07;
 DT 01-DEC-2001 (TfEMBLrel. 19, Created)
 DT 01-DEC-2001 (TfEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TfEMBLrel. 20, Last annotation update)
 DE MDC-SIGNIA type III isoform.
 GN CD209.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=1137487;
 RA Mummidi S., Catano G., Lam L., Hoelle A., Telles V., Begun K.,
 RJ Jimenez F., Ahuja S.S., Ahuja S.K.;
 RT *Extensive repertoire of Membrane-bound and Soluble pendritic Cell-
 specific ICAm-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
 RT Isoforms. Inter-individual Variation in Expression of DC-SIGN
 RT Transcripts.*;
 RL J. Biol. Chem. 276:33196-33212(2001).
 DR EMBL: AY042323; AAK91848.1;
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF00059; Lectin_C.1;
 DR PROSITE: PS00615; C-TYPE_LECTIN_1; UNKNOWN_1.
 DR PROSITE: PS00641; C-TYPE_LECTIN_2; 1;
 SO SEQUENCE 312 AA; 35178 MW; 6576161AF64ABE CRC64;

Query Match 24.0%; Score 78; DB 4; Length 412;
 Best Local Similarity 100.0%; Pred No. 3.5e-72;
 Matches 78; Conservative 0; Mismatches 0; Indels 0;

QY 58 QDAIYONLTOLKAAGVGLSKLOEIYQELTOLKAAGVGLPEKSKLOEIYQELTIRLAKA 117
 DB 74 QDAIYONLTOLKAAGVGLSKLOEIYQELTOLKAAGVGLPEKSKLOEIYQELTIRLAKA 109
 QY 118 VGLPEKSKLOEIYQELT 135
 DB 134 VGLPEKSKLOEIYQELT 127

RESULT 14

Q96Q99 PRELIMINARY; PRT; 334 AA.
 ID Q96Q99;
 AC Q96Q99;
 DT 01-DEC-2001 (TfEMBLrel. 19, Created)
 DT 01-DEC-2001 (TfEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TfEMBLrel. 20, Last annotation update)
 DE SDC-SIGNIB type II isoform.
 GN CD209.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=1137487;
 RA Mummidi S., Catano G., Lam L., Hoelle A., Telles V., Begun K.,
 RJ Jimenez F., Ahuja S.S., Ahuja S.K.;
 RT *Extensive Repertoire of Membrane-bound and Soluble pendritic Cell-
 specific ICAm-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
 RT Isoforms. Inter-individual Variation in Expression of DC-SIGN
 RT Transcripts.*;
 RL J. Biol. Chem. 276:33196-33212(2001).
 DR EMBL: AY042331; AAK91856.1;
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF00059; Lectin_C.1;
 DR PROSITE: PS00615; C-TYPE_LECTIN_1; UNKNOWN_1.
 DR PROSITE: PS00641; C-TYPE_LECTIN_2; 1;
 SO SEQUENCE 334 AA; 37843 MW; BE766FC4111CB6BF CRC64;

```

Query Match: 24.0%; Score 78; DB 4; Length 334;
Best Local Similarity 100.0%; Pred. No. 3 7e-72;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QDAIYQNI.TOLKAAVGLSEKSKLOEYQETIQLKAAVGLPEKSKLOEYQETRLKAA 117
    |||
DB 50 QDAIYQNI.TOLKAAVGLSEKSKLOEYQETIQLKAAVGLPEKSKLOEYQETRLKAA 109

QY 118 VGLPEKSKLOEYQET 135
    |||
DB 110 VGLPEKSKLOEYQET 127

RESULT 15
Q96004 PRELIMINARY PRT: 360 AA.
ID Q96004
AC Q96004
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DE 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE SPC-SIGNIA type II isoform.
CN C0209.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
NX NCBI_TaxID=9606;
RN 11
RF SEQUENCE FROM N.A.
RX PubMed=11337487;
RA Mummidi S., Carano G., Lam L., Hoeftle A., Telles V., Begum K.,
KA Jimenez F., Ahuja S.S., Ahuja S.K.;
KT "Extensive Repertoire of Membrane-bound and Soluble Benthic Cell-
RT specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
RT Isoforms: Inter-individual Variation in Expression of DC-SIGN
RT Transcripts."
PL J. Biol. Chem. 276:33196-33212(2001).
DR EMBL: AY042226; AAK91851.1;
DR InterPro: IPR001304; Lectin_C;
DR Pfam: PF00059; Lectin_C;
DR PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE: PS00615; C_TYPE_LECTIN_2;
SQ SEQUENCE 360 AA: 41009 MW: 648238AARDFAAAH CRC64;

Query Match: 24.0%; Score 78; DB 4; Length 360;
Best Local Similarity 100.0%; Pred. No. 4e-72;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QDAIYQNI.TOLKAAVGLSEKSKLOEYQETIQLKAAVGLPEKSKLOEYQETRLKAA 117
    |||
DB 30 QDAIYQNI.TOLKAAVGLSEKSKLOEYQETIQLKAAVGLPEKSKLOEYQETRLKAA 89

QY 118 VGLPEKSKLOEYQET 135
    |||
DB 90 VGLPEKSKLOEYQET 107

RESULT 16
Q96005 PRELIMINARY PRT: 380 AA.
ID Q96005
AC Q96005
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DE 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE SPC-SIGNIA type I isoform.
CN C0209.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
NX NCBI_TaxID=9606;
RN 11
RF SEQUENCE FROM N.A.
RX PubMed=11337487;

```

```

RA Mummidi S., Carano G., Lam L., Hoeftle A., Telles V., Begum K.,
KA Jimenez F., Ahuja S.S., Ahuja S.K.;
KT "Extensive Repertoire of Membrane-bound and Soluble Benthic Cell-
RT specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
RT Isoforms: Inter-individual Variation in Expression of DC-SIGN
RT Transcripts."
PL J. Biol. Chem. 276:33196-33212(2001).
DR EMBL: AY042226; AAK91851.1;
DR InterPro: IPR001304; Lectin_C;
DR Pfam: PF00059; Lectin_C;
DR PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE: PS00615; C_TYPE_LECTIN_2;
SQ SEQUENCE 360 AA: 43330 MW: BCFWCC45AEFE6B02 CRC64;

Query Match: 24.0%; Score 78; DB 4; Length 380;
Best Local Similarity 100.0%; Pred. No. 4.2e-72;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QDAIYQNI.TOLKAAVGLSEKSKLOEYQETIQLKAAVGLPEKSKLOEYQETRLKAA 117
    |||
DB 50 QDAIYQNI.TOLKAAVGLSEKSKLOEYQETIQLKAAVGLPEKSKLOEYQETRLKAA 109

QY 118 VGLPEKSKLOEYQET 135
    |||
DB 110 VGLPEKSKLOEYQET 127

RESULT 17
Q96000 PRELIMINARY PRT: 380 AA.
ID Q96000
AC Q96000
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DE 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE SPC-SIGNIA type I isoform.
CN C0209.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
NX NCBI_TaxID=9606;
RN 11
RF SEQUENCE FROM N.A.
RX PubMed=11337487;
RA Mummidi S., Carano G., Lam L., Hoeftle A., Telles V., Begum K.,
KA Jimenez F., Ahuja S.S., Ahuja S.K.;
KT "Extensive Repertoire of Membrane-bound and Soluble Benthic Cell-
RT specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
RT Isoforms: Inter-individual Variation in Expression of DC-SIGN
RT Transcripts."
PL J. Biol. Chem. 276:33196-33212(2001).
DR EMBL: AY042226; AAK91851.1;
DR InterPro: IPR001304; Lectin_C;
DR Pfam: PF00059; Lectin_C;
DR PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE: PS00615; C_TYPE_LECTIN_2;
SQ SEQUENCE 380 AA: 43125 MW: F3D096F9F7D044B CRC64;

Query Match: 24.0%; Score 78; DB 4; Length 380;
Best Local Similarity 100.0%; Pred. No. 4.2e-72;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QDAIYQNI.TOLKAAVGLSEKSKLOEYQETIQLKAAVGLPEKSKLOEYQETRLKAA 117
    |||
DB 50 QDAIYQNI.TOLKAAVGLSEKSKLOEYQETIQLKAAVGLPEKSKLOEYQETRLKAA 109

QY 118 VGLPEKSKLOEYQET 135
    |||
DB 110 VGLPEKSKLOEYQET 127

RESULT 18
Q96008 PRELIMINARY PRT: 398 AA.
ID Q96008

```

AC 096008: 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DE MDC-SIGNIA type II isoform.
 GN CD209.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed:11337487;
 RA Mummidi S., Catano G., Lam L., Hoeft A., Telles V., Begum K.,
 RA Jimenez F., Ahuja S.S., Ahuja S.K.;
 RT "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
 RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
 RT Isoforms Inter-individual Variation in Expression of DC-SIGN
 RT Transcripts";
 RL J. Biol. Chem. 276:33196-33212(2001).
 DR EMBL: AY042222; AAK91847.1;
 DR InterPro: IPR001304; Lectin_C;
 DR Pfam: PF00059; Lectin_C; 1;
 DR PROSITE: PS00615; C-TYPE_LLECTIN_1; UNKNOWN_1.
 DR PROSITE: PS50041; C-TYPE_LLECTIN_2; 1;
 SQ SEQUENCE 398 AA: 45031 MW: D9257648679123E CRC64;

Query Match 24.08: Score 78; DB 4; Length 398;
 Best Local Similarity 100.0%; Pred. No. 4, 4e-72;

Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QDAIYONITOLKAAGVSEKSLQEIYQELTOLKAAGVSEKSLQEIYQELTRKAA 117
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 74 QDAIYONITOLKAAGVSEKSLQEIYQELTOLKAAGVSEKSLQEIYQELTRKAA 133
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 118 VGELEPKSKLOEIYQELT 135
 ||||||||||||||||||||||||
 DB 134 VGELEPKSKLOEIYQELT 151

RESULT 19

Q96001 PRELIMINARY; PRT; 404 AA.

AC 096001: 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DE MDC-SIGNIB type I isoform.
 GN CD209.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed:11337487;
 RA Mummidi S., Catano G., Lam L., Hoeft A., Telles V., Begum K.,
 RA Jimenez F., Ahuja S.S., Ahuja S.K.;
 RT "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
 RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
 RT Isoforms Inter-individual Variation in Expression of DC-SIGN
 RT Transcripts";
 RL J. Biol. Chem. 276:33196-33212(2001).
 DR EMBL: AY042222; AAK91847.1;
 DR InterPro: IPR001304; Lectin_C;
 DR Pfam: PF00059; Lectin_C; 1;
 DR PROSITE: PS00615; C-TYPE_LLECTIN_1; UNKNOWN_1.
 DR PROSITE: PS50041; C-TYPE_LLECTIN_2; 1;
 SQ SEQUENCE 404 AA: 45570 MW: E4D570727D9DC330 CRC64;

Query Match 24.08: Score 78; DB 4; Length 404;
 Best Local Similarity 100.0%; Pred. No. 4, 4e-72;
 Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QDAIYONITOLKAAGVSEKSLQEIYQELTOLKAAGVSEKSLQEIYQELTRKAA 117
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 74 QDAIYONITOLKAAGVSEKSLQEIYQELTOLKAAGVSEKSLQEIYQELTRKAA 133
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 118 VGELEPKSKLOEIYQELT 135
 ||||||||||||||||||||||||
 DB 134 VGELEPKSKLOEIYQELT 151

RESULT 20

Q9NNX6 PRELIMINARY; PRT; 404 AA.

AC 09NNX6: 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Membrane-associated lectin type-2 (probable mannose-binding
 DE lectin DC-SIGN) (MDC-SIGNIA type I isoform).
 GN CD209.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed:11337487;
 RA Mummidi S., Catano G., Lam L., Hoeft A., Telles V., Begum K.,
 RA Jimenez F., Ahuja S.S., Ahuja S.K.;
 RT "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
 RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
 RT Isoforms Inter-individual Variation in Expression of DC-SIGN
 RT Transcripts";
 RL J. Biol. Chem. 276:33196-33212(2001).
 DR EMBL: M98457; AAF77072.1;
 DR EMBL: AF209479; AAG13814.1;
 DR EMBL: AF209486; AAK20997.1;
 DR EMBL: AY042221; AAK91846.1;
 DR HSP: P22897; IEGG;
 DR InterPro: IPR001304; Lectin_C;
 DR Pfam: PF00059; Lectin_C; 1;
 DR SMART: SM00044; CLECT; 1;
 DR PROSITE: PS00615; C-TYPE_LLECTIN_1; UNKNOWN_1.
 DR PROSITE: PS50041; C-TYPE_LLECTIN_2; 1;
 SQ SEQUENCE 404 AA: 45775 MW: A25FA24601453360 CRC64;

FT	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
FT	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
FT	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
FT	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
FT	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
FT	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80																				


```

OX  NCB1_taxID=9544;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=21819420; PubMed=11818554;
RA  Wu L., Bashirova A.A., Martin T.D., Villamide L., Mehlopp E.,
RT  "Rhesus macaque dendritic cells efficiently transmit primate
RL  lentiviruses independently of DC-SIGN."
RP  Proc. Natl. Acad. Sci. U.S.A. 99:1508-1514(2002).
DP  EMBL: AY040319; AAK74185.1; -.
SQ  SEQUENCE 381 AA; 42955 MW; 073E7B0B42C91049 CRC64;

Query Match 12.3%; Score 40; DB 6; Length 381.
Best Local Similarity 100.0%; Pred. No. 9, 9e-33.
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  110 ELTRKAAVGLPEPSKLOEIYQELTRKAAVGLPEPSK 149
DB  126 ELTRKAAVGLPEPSKLOEIYQELTRKAAVGLPEPSK 165
|||||
RESULT 25
QY5JC6 PRELIMINARY; PRT; 381 AA.
ID QY5JC6;
AC QY5JC6;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Dendritic cell-specific ICAM-3 grabbing nonintegrin.
OS Macaca nemestrina (pig-tailed macaque).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciinae;
OC Cercopitheciinae; Macaca.
NX NCBI_TaxID=9545;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21465051; PubMed 11581396;
RA Haribaud F., Pohlmann S., Sparwasser T., Kimura M.T., Choi Y.-C.,
RA Harigarty B.S., Ahmad N., Macfarlan T.G., Leslie G.J.,
RA Arnsperger C., Reinhardt T.A., Kimata J.T., Hoffman D.R., Hoxle J.A.,
RA Doms R.W.;
RT "Functional and antigenic characterization of human, rhesus macaque,
RT pig-tailed macaque, and murine dc-sigm."
RL J. Virol. 75:10281-10289(2001).
DR EMBL: AF343727; AAL14428.1; -.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C.
DR PROSITE: PS00615; C-TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
DR Integrin.
RX NCBI_TaxID=9546;
RN [1]
RP SEQUENCE 381 AA; 42951 MW; C4F6E2F1D454B74A CRC64;

Query Match 10.5%; Score 34; DB 6; Length 381.
Best Local Similarity 100.0%; Pred. No. 1, 6e-26.
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  93 AAVGELPEPSKLOEIYQELTRKAAVGLPEPSK 126
DB  132 AAVGELPEPSKLOEIYQELTRKAAVGLPEPSK 165
|||||
RESULT 26
QY5J96 PRELIMINARY; PRT; 404 AA.
ID QY5J96;
AC QY5J96;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Dendritic cell-specific ICAM-3 grabbing non-integrin.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciinae;

```

```

OC Cercopitheciinae; Macaca.
OX NCB1_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=LYMPH NODE;
RA Gajtenieck T.R.H., Koopman G., van Duijnhoven S.C.P., van
RA van Schijndel A.C., Engering A., Heeney J.L., van Kooyk
RT "Identification of Rhesus and Chimpanzee DC-SIGN, both a
RT gp120 trans-receptors similar as human DC-SIGN."
RL Submitted (JUN-2001) to the EMBL/GenBank/TrEMBL databases.
DR EMBL: AF310886; AAK97459.1; -.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C.
DR PROSITE: PS00615; C-TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
RX Integrin.
SQ SEQUENCE 404 AA; 45655 MW; E7753413430CC17 CRC64;

Query Match 8.0%; Score 26; DB 6; Length 404.
Best Local Similarity 100.0%; Pred. No. 3, 4e-18.
Matches 26; Conservative 0; Mismatches 0; Indels 0;

QY  110 ELTRKAAVGLPEPSKLOEIYQEL 145
DB  126 ELTRKAAVGLPEPSKLOEIYQEL 151
|||||
RESULT 27
QY6Q06 PRELIMINARY; PRT; 168 AA.
ID QY6Q06;
AC QY6Q06;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE MDC-SIGNA type IV isoform.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=1137487;
RA Mummidi S., Catano G., Lam L., Hoxle A., Telles V., Bee
RA Jimenez F., Ahuja S.S., Ahuja S.K.;
RT "Extensive repertoire of Membrane-bound and soluble lectin
RT specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN) and its
RT isoforms: Inter-individual Variation in Expression of Dc
RT transcripts."
RL J. Biol. Chem. 276:33196-33212(2001).
DR EMBL: AY042224; AAK91849.1; -.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C.
DR PROSITE: PS00615; C-TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
SQ SEQUENCE 168 AA; 18646 MW; 9017BC79C489F59C CRC64;

Query Match 6.5%; Score 21; DB 4; Length 168.
Best Local Similarity 100.0%; Pred. No. 2, 4e-13.
Matches 21; Conservative 0; Mismatches 0; Indels 0;

QY  253 WAGSLDNOEGTQWQVDSPL 273
DB  79 WAGSLDNOEGTQWQVDSPL 99
|||||
RESULT 28
QYV1K4 PRELIMINARY; PRT; 207 AA.
ID QYV1K4;
AC QYV1K4;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)

```

DE SIGNR3.
 GH CD209D OR SIGNR3.
 OS Mus musculus (Mouse).
 AC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 GN NCB1_TaxID=10090;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RZ Park G.G., Steinman R.M.;
 R1 "Alternatively Spliced Forms of Mouse DC-SIGN Homologs."
 R1 Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RF EMBL: AF440280; AAL33584.1;
 DE MGI: MGI:2157947; CD209d.
 DB MGI: MGI:2157947; CD209d.
 DE InterPro: IPR001304; Lectin_C.
 DE Pfam: PF00059; Lectin_C; 1.
 DE SMART: SM00034; CLECT_1
 DE PROSITE: PS00615; C-TYPE_LECTIN_1; UNKNOWN_1.
 DE PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
 SC SEQUENCE 207 AA, 23500 MW, 16744.477DAAGSR CWC64.

Query Match 3.48; Score 11; DB 11; Length 207;
 Best Local Similarity 100.0%; Pred No. 0.0065;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 WDCSPLSPSF 277
 DB 148 WDCSPLSPSF 158
 |||||

RESULT 29
 Q91ZW8 PRELIMINARY; PRT; 237 AA
 ID Q91ZW8
 AC Q91ZW8
 DT 01-DEC-2001 (TEMBLrel. 19, Created)
 DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
 DE SIGNR3.
 GN CD209D OR SIGNR3.
 OS Mus musculus (Mouse).
 AC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 GN NCB1_TaxID=10090;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RZ Park G.G., Takahara K., Umemoto E., Yashima Y., Matsubara K.,
 RA Matsuda Y., Clausen B.E., Inaba K., Steinman R.M.;
 RA "Five mouse homologues of the human dendritic cell C-type lectin, DC-SIGN."
 RT SIGN."
 RI 1st. Immunol. 13:1283-1290(2001).
 DE EMBL: AF373411; AAL1237.1;
 DE MGI: MGI:2157947; CD209d
 DE InterPro: IPR001304; Lectin_C
 DE Pfam: PF00059; Lectin_C; 1.
 DE PROSITE: PS00615; C-TYPE_LECTIN_1; UNKNOWN_1.
 DE PROSITE: PS50041; C-TYPE_LECTIN_2; 1
 SC SEQUENCE 237 AA, 26925 MW, 52830.925CAF8HFC CWC64;

Query Match 3.48; Score 11; DB 11; Length 217;
 Best Local Similarity 100.0%; Pred No. 0.0073;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 WDCSPLSPSF 277
 DB 178 WDCSPLSPSF 188
 |||||

RESULT 30
 Q16641 PRELIMINARY; PRT; 168 AA
 AC Q16641
 Q16641

DT 01-JAN-1998 (TEMBLrel. 05, Created)
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
 DE Epidermal 18.2 KDa protein.
 GN W09610.5.
 OS Caenorhabditis elegans.
 AC Fokoyeta, M., Lucas K.M., O'Keefe M.A., Hochrein H., Laabi Y.,
 OC Rhabditidae; Polychaeta; Caenorhabditis.
 OC Rhabditidae; Polychaeta; Caenorhabditis.
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE: 99059613; PubMed: 9851916;
 RA Nere;
 RT "Genome sequence of the nematode *C. elegans*: a platform for
 RT Investigating biology. The *C. elegans* Sequencing Consortium."
 RL Science 282:2012-2018(1998).
 RN 121
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Bentley D., Goebl D., Holmes A.;
 RT "The sequence of *C. elegans* cosmid W09610."
 RL Submitted (AUG 1997) to the EMBL/GenBank/DBJ databases.
 RN 131
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission."
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF016671; AAB66113.1;
 DR HSSP: P22897; IREG.
 DR InterPro: IPR001304; Lectin_C.
 DE Pfam: PF00059; Lectin_C; 1.
 DE SMART: SM00034; CLECT_1
 DE PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
 DR Hypothetical protein.
 SC SEQUENCE 108 AA; 18217 MW; 494FE3ADU031BRC5 CWC64;

Query Match 2.88; Score 9; DB 5; Length 168;
 Best Local Similarity 100.0%; Pred No. 0.64;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 YWNSCEPNN 288
 DB 123 YWNSCEPNN 131
 |||||

RESULT 31
 Q91ZX1 PRELIMINARY; PRT; 238 AA
 ID Q91ZX1
 AC Q91ZX1
 DT 01 DEC 2001 (TEMBLrel. 19, Created)
 DT 01 DEC 2001 (TEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
 DE DC-SIGN (Dendritic cell C-type lectin protein CIRE).
 GN CD209A OR DC-SIGN.
 OS Mus musculus (Mouse).
 AC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 GN NCB1_TaxID=10090;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RZ Park G.G., Takahara K., Umemoto E., Yashima Y., Matsubara K.,
 RA Matsuda Y., Clausen B.E., Inaba K., Steinman R.M.;
 RA "Five mouse homologues of the human dendritic cell C-type lectin, DC-SIGN."
 RT SIGN."
 RI 1st. Immunol. 13:1283-1290(2001).
 DE EMBL: AF373411; AAL1237.1;
 DE MGI: MGI:2157947; CD209d
 DE InterPro: IPR001304; Lectin_C
 DE Pfam: PF00059; Lectin_C; 1.
 DE PROSITE: PS00615; C-TYPE_LECTIN_1; UNKNOWN_1.
 DE PROSITE: PS50041; C-TYPE_LECTIN_2; 1
 SC SEQUENCE 237 AA, 26925 MW, 52830.925CAF8HFC CWC64;

```

KA Hirotsuki T.C., Lew A.M., Shortman K., Wright M.D.:
RT "Molecular cloning of a C-type lectin superfamily protein
RT differentially expressed by CD8alpha(-) splenic dendritic cells."
RL Mol. Immunol. 38:365-373(2001).
DR EMBL: AF373408; AAL13234.1; .
DR EMBL: AY049062; AAK85825.1; .
DR MGI: 2157943; Cd209a.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C_1.
DR PROSITE: PS00615; C_TYPE_LECTIN_1; UNKN/MN_1.
DR PROSITE: PS50041; C_TYPE_LECTIN_2; 1.
KW Lectin.
SQ SEQUENCE 238 AA: 27149 MW: 95025 PFC2E6DC2 CRG64;

Query Match 2.88; Score 9; DB 11; Length 238;
Best Local Similarity 100.0%; Pred. No. 0.87;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 YQELTOLKA 93
DB 91 YQELTOLKA 99
|||||

RESULT 32
Q8TB07 PRELIMINARY: PKT: 491 AA.
ID Q8TB07:
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DR Hypothetical 53.0 kDa protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
ON 11
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE=UTERUS;
RA Strausberg R.;
RL Submitted (FEF8-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC024214; AAI24214.1; .
KW Hypothetical protein.
FT NON_TER
SU SEQUENCE 491 AA: 53045 MW: 53043838496A08 CRG64;

Query Match 2.88; Score 9; DB 4; Length 491;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 AGVIVAILV 42
DB 278 AGVIVAILV 286
|||||

RESULT 33
Q96002 PRELIMINARY: PKT: 34 AA.
ID Q96002:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DR SDC-STIGNIA type IV isoform.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
ON 11
RN 11
RP SEQUENCE FROM N.A.
RC PubMed=11337487;
RA Jimenez F., Catano G., Lam L., Hocile A., Telles V., Begum K.,
RA Jimenez F., Ahuja S.S., Ahuja S.K.;
RT "Extensive Repertoire of Membrane-Bound and Soluble Dendritic Cell-
```

```

RT Specific 10AM-3-mabbing Nucleoside 1 (Nucleoside) and 1st
RT Isoforms. Inter-individual Variation in Expression of 1st
RT Transcripts."
RL J. Biol. Chem. 276:33196-33212(2001).
DR FMBL: AY042228; AAK91853.1; .
SQ SEQUENCE 34 AA: 4046 MW: 5155A69111A0150 CRG64;

Query Match 2.59; Score 8; DB 4; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY 1 MSDSKPKR 8
DB 1 MSDSKPKR 8
|||||

RESULT 34
Q9LFZ0 PRELIMINARY: PKT: 88 AA.
ID Q9LFZ0:
AC Q9LFZ0:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE T7N9.1
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tr
OC Spermatophyta; Magnoliophyta; Eudicotyledons; Core eudi
OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
ON 11
RN 11
RP SEQUENCE FROM N.A.
RA Khan S., Brooks S., Buchler E., Chao G., Johnson-Hopson
RA Shinn P., Altali H., Bei W., Chin C., Chou J., Choi E.,
RA Conway A., Gonzales A., Hansen N., Howing H.,
RA Lenz G., Li J., Liu A., Liu J., Liu S., Makharas
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Toriumi M., Vaysberg M., Yu G., Fiederspiel N.A., Theodor
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC T7N9.1"
RT Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases
KL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases
RN 12
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases
RN 13
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases
RN 14
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases
RN 15
RP SEQUENCE FROM N.A.
RA Cheuk P., Shinn P., Brooks S., Buchler E., Chao G., Joh
RA Khan S., Kim C., Altali H., Bei W., Chin C., Chou J.,
RA Conn J., Conway A., Gonzalez A., Hansen N., Howing H.,
RA Lee J., Lenz G., Li J., Liu A., Liu J., Liu S., Makharas
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., S
RA Thaler A., Toriumi M., Vaysberg M., Yu G., Davis R.,
RA Theologis A., Ecker J.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases
RN 16
RP SEQUENCE FROM N.A.
RA Cheuk P., Shinn P., Brooks S., Buchler E., Chao G., Joh
RA Khan S., Kim C., Altali H., Bei W., Chin C., Chou J.,
RA Conn J., Conway A., Gonzalez A., Hansen N., Howing H.,
RA Lee J., Lenz G., Li J., Liu A., Liu J., Liu S., Makharas
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., S
RA Thaler A., Toriumi M., Vaysberg M., Yu G., Davis R.,
RA Theologis A., Ecker J.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases
```

```

D> EMBL: AC000348; AAF79848.1;
SQ SEQUENCE 88 AA; 10203 MW; 494014941809444 CRC64;
Query Match:
  Best Local Similarity 100.0%; Score 8, DB 10, Length 88;
  Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Q> 25 L01LSEFM.32
  11111111
Db 38 L01LSEFM.45

RESULT 35
Q91ZW7 PRELIMINARY; PRT: 298 AA.
ID Q91ZW7;
AC Q91ZW7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DI 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
DI 01-JUN-2002 (TrEMBLrel. 21, last annotation update)
DE SIGNR4.
GN C02094; OR SIGNR4.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
PA 11;
RE SEQUENCE FROM N A
RC STRAIN=C57BL/6;
RX PubMed-11581173;
RA Park G.G., Takahara K., Umecoro E., Yashima Y., Matsubara K.,
RA Matsuda Y., Clausen R.E., Inaba K., Steinman R.M.;
RT "Five mouse homologues of the human dendritic cell C-type lectin, DC-
RT SIGN-4";
RL Int. Immunol. 13:1283-1290(2001).
DR EMBL: AF373412; AAL13238.1;
DR MGD: MGI:2157948; Cd209c.
DR InterPro: IPR001304; Lectin_C
DR Pfam: PF00059; Lectin_C; 1.
DR PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE: PS00641; C_TYPE_LECTIN_2; 1.
SV SEQUENCE 298 AA; 24382 MW; 219471E0870DEB7 CRC64.

Query Match:
  Best Local Similarity 100.0%; Score 8, DB 11, Length 208.
  Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Q> 253 WMG1SDLN.260
  11111111
Db 135 WMG1SDLN.142

RESULT 36
Q94ZD4 PRELIMINARY; PRT: 279 AA.
ID Q94ZD4;
AC Q94ZD4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DI 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
DI 01-JUN-2002 (TrEMBLrel. 21, last annotation update)
DE A1912410/FS011_7.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC eustoids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
PA 11;
RE SEQUENCE FROM N.A.
RC CHOU P., CHEN H., KIM C.T., KOSOMA F., MEYERS M.C., BARTH T.,
RA BOWSER L., CARLINI P., DALE J.M., GILDSMITH A.D., HAYASHIZAKI Y.,
RA ISHIDA J., JIANG P.X., JONES T., KAMIYA A., KARLIN-NEUMANN G.,
RA KAWAI J., LAM B., LEE J.M., LIN J., LIU S.X., MIRANDA M., NARUSAKA M.,
RA NGUYEN M., ONODERA C.S., PALM C.J., PHAM P.K., QUACH H.L., SAKURAI T.,
RA SATOU M., SEKI M., SOUTHWICK A., TANG C.C., TORIUMI M., YAMADA K.,
RA YAMAMURA Y., YU G., YU S., SHINOZAKI K., DAVIS P.W., THEOLOGIS A.,

```

```

RA ECKER J.R.;
RT "Arabidopsis cDNA clones";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY057617; AAL44412.1;
DR MEROPS: S14.001;
DR InterPro: IPR001907; CLP_protease.
DR Pfam: PF00574; CLP_protease; 1.
DR TrEMBL: F7FFD2C3374830B CRC64.
SQ SEQUENCE 279 AA; 31150 MW; F7FFD2C3374830B CRC64.

Query Match:
  Best Local Similarity 100.0%; Score 8, DB 10, Length 279;
  Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Q> 262 EGTWQWD.269
  11111111
Db 82 EGTWQWD.89

RESULT 37
Q9XJ36 PRELIMINARY; PRT: 279 AA.
ID Q9XJ36;
AC Q9XJ36;
DI 01-NOV-1999 (TrEMBLrel. 12, Created)
DI 01-NOV-1999 (TrEMBLrel. 12, last sequence update)
DI 01-JUN-2002 (TrEMBLrel. 21, last annotation update)
DE NCLP2 (FS011.13) (Similar to nclp2 db|BA082066.1).
GN NCLP2; OR ATG12410, FS011.13 OR ATG12410.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC eustoids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
PA 11;
RE SEQUENCE FROM N.A.
RC STRAIN=COLOMBIA;
RA NARAYANASHY K., FRE M., KINGSLEY T., SHINOZAKI K., WATANABE A.;
RT "Identification of clp genes expressed in senescing Arabidopsis
RT leaves";
RL Plant Cell Physiol. 0:0-0(1999).
SV 13;
EN 11;
RE SEQUENCE FROM N.A.
RC CHOU P., BROOKS S., BUEHLER E., JOHNSON-HOPSON C., KHAN S., KIM C.,
RA SHIN P., ALLALI H., BEL G., CHIN C., CHOU J., CHOI E., CONN L.,
RA CONWAY A., GONZALES A., HANSEN R., HOWING B., KOO T., LAM B., LEE J.,
RA LENZ G., LI J., LIU A., LIU K., LIU S., MUKHARSKY N., NGUYEN M.,
RA PALM C., PHAM P., SAKANO H., SCHWARTZ J., SOUTHWICK A., THAYER A.,
RA TORIUMI M., VAYSBERG M., YU G., FEDERSPIEL N.A., THEOLOGIS A.,
RA ECKER J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC FS011 from chromosome
RT 1";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RE SEQUENCE FROM N.A.
RC ECKER J.R.;
RA Submitted (MAR 2000) to the EMBL/GenBank/DBJ databases.
RT ECKER J.R.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
SV 15;
RE SEQUENCE FROM N.A.
RC CHOU P., SHIN P., BROOKS S., BUEHLER E., CHAO Q., JOHNSON-HOPSON C.,
RA KHAN S., KIM C., ALLALI H., BEL G., CHIN C., CHOU J., CHOI E.,
RA CONN L., CONWAY A., GONZALES A., HANSEN R., HOWING B., KOO T., LAM B.,
RA LEE J., LEE G., LI J., LIU A., LIU J., LIU S., MUKHARSKY N.,
RA NGUYEN M., PALM C., PHAM P., SAKANO H., SCHWARTZ J., SOUTHWICK A.,
RA THAYER A., TORIUMI M., VAYSBERG M., YU G., DAVIS R., FEDERSPIEL N.,
RA THEOLOGIS A., ECKER J.;
PL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases
SV 16;
RE SEQUENCE FROM N.A.
RC SOUTHWICK A., KARLIN-NEUMANN G., NGUYEN M., LAM B., MIRANDA M.,

```

RA Palm C.J., Bowser L., Jones T., Bant J., Carinci P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shin P., Yamada K., Shinozaki K.,
 RA Ecker J., Theologis A., Davis R.W.,
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RA Southwick A., Karlin-Neumann G., Nguyen M., Lam H., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Bant J., Carinci P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shin P., Yamada K., Shinozaki K.,
 RA Ecker J., Theologis A., Davis R.W.,
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB022327; HAAB2066.1;
 DR EMBL: AC025416; AAF79635.1;
 DR EMBL: AY062770; AAL32848.1;
 DR EMBL: AY081641; AAM10203.1;
 DR HSSP: P19245; 1TF;
 DR MEROPS: S14.001;
 DR InterPro: IPR001907; CLP-Protase;
 DR Pfam: PF00574; CLP-Protase; 1;
 DR PRINTS: PR00127; CLP-PROTEASE;
 DR Trifams: TIGR00494; CLP: 1;
 SU SEQUENCE 279 AA; 31208 MW; DADICFDND478408 CPG64;

Query Match 2.58; Score 8; DB 10; Length 279;
 Best Local Similarity 100.0%; Prod. No. 11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 ECTMOWND 269
 |||||
 DB 42 ECTMOWND 89

RESULT 38
 ID P72781 PRELIMINARY; PFI: 282 AA.
 AC P72781;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Regulatory components of sensory transduction system.
 GN SLR1783.
 OS Synecocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis
 CX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE:97061201; PubMed 8905231.
 RA Kureko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hirosewa M., Saito M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., of jmurra S.,
 RA Shimpou S., Takeuchi C., Wada T., Motanabe A., Yamada M., Yasuda M.,
 RA Taira S.,
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synecocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions."
 RL DNA Res. 3:109-146(1996)
 CC -1- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
 CC REGULATORY COMPONENTS OF SENSOR TRANSDUCTION SYSTEMS.
 CC -1- SIMILARITY: BELONGS TO THE LUXR/DHHA FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 DR EMBL: D90900; BAAL6796.1;
 DR HSSP: P08402; 1B00.
 DR InterPro: IPR000792; HTH_LuxR.
 DR InterPro: IPR001789; Response_Treg.
 DR Pfam: PF00196; GORE; 1;
 DR Pfam: PF00072; response_Treg; 1;
 DR Prodom: PD000039; response_Treg; 1;
 DR Prodom: PD000307; HTH_LuxR; 1;
 DR SMART: SM00421; HTH_LuxR; 1;

DR SMART: SM00448; REC; 1;
 KW DNA binding; Phosphorylation; Sensory transduction;
 KW Transcription regulation; Complete proteome;
 SU SEQUENCE 282 AA; 31395 MW; IAF76614B766B4B7 CPG64;

Query Match 2.58; Score 8; DB 10; Length 282;
 Best Local Similarity 100.0%; Prod. No. 11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY 149 KLOEYOE 156
 |||||
 DB 149 KLOEYOE 156

RESULT 39
 ID Q9DBV4 PRELIMINARY; PFI: 41 AA.
 AC Q9DBV4;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE 1810030122RIK protein.
 CN CD2098 OR 1810040122RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Cladoda; Vertebrata; Eute-
 OC Mammalia; Eutheria; Rodentia; Sturnoethia; Muridae; Mur
 CX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN:G57H/6J; TISSUE:PANCREAS;
 RX MEDLINE=21085660; PubMed=1121781;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yae
 RA Saito T., Okazaki Y., Gojohori I., Hone H., Kasukawa Y.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casady
 RA Fleischmann W., Gaasterland T., Gissi G., King R., Koci
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G.,
 RA Schindl L., Staudt F., Suzuki K., Tamita M., Wagner J.,
 RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., de
 RA Blake J., Boffelli D., Bojunga N., Carinci P., de Heide
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Garib
 RA Garsuch S., Hill D., Holtman M., Hume D.A., Kamiya M.
 RA Lyons P., Marchionni L., Mashima T., Mazzaroli J., Mont
 RA Nordone P., Rind B., Ringwald M., Rodriguez I., Sakamot
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y.,
 RA Suzuki H., Toyooka K., Weng K., Weller G., Whitaker
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai J., K
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA colle
 RL Nature 409:985-990(2001).
 DR EMBL: AK087656; HA025165.1;
 DR HSSP: P22897; 1P6G.
 DR M50: M51:1916415; GAD209B.
 DR InterPro: IPR001353; AntiProtein-1.
 DR InterPro: IPR001304; Lectin_C1.
 DR Pfam: PF00059; Lectin_C1;
 DR PRINTS: PR00356; ANTIPEPTEIN.
 DR SMART: SM00034; CLECT; 1;
 DR PROSITE: PS00615; G_TYPE_LECTIN_1; UNKNOWN_1;
 DR PROSITE: PS50041; G_TYPE_LECTIN_2; 1;
 SU SEQUENCE 311 AA; 35618 MW; 923550302EF9241 CPG64;

Query Match 2.58; Score 8; DB 10; Length 30;
 Best Local Similarity 100.0%; Prod. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY 23 LVLOLSF 30
 |||||
 DB 40 LVLOLSF 47

RESULT 40

```

0912X0
ID 0912X0 PRELIMINARY: PRT: 325 AA.
AC 0912X0:
DT 01-DEC-2001 (TREMblrel, 19, Created)
DT 01-DEC-2001 (TREMblrel, 19, Last sequence update)
DE 01-JUN-2002 (TREMblrel, 21, Last annotation update)
DE SIGNR1.
GN CD209H OR SIGNR1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN 11
RC SEQUENCE FROM N A
RC STRAIN=C57BL/6;
RX PubMed=11581173;
KA Park G.G., Takahara K., Imamoto F., Yashima Y., Matsubara K.,
KA Matsuda Y., Clausen B.F., Inaba K., Steinman P.M.;
R1 "Five mouse homologues of the human dendritic cell C-type lectin, DC
R1 SIGN.";
R1 Int. Immunol. 13:1283-1290(2001).
DR EMBL: AE373409; AAL13235.1; -.
DR MGD: MGI:1916415; G62998.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C.1.
DR PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE: PS00641; C_TYPE_LECTIN_2; 1.
SQ SEQUENCE 325 AA: 37111 MW: 96938640/24704 CPG64;

Query Match 2.5%; Score 8; DB 11; Length 325;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 23 LVLGLSEF 30
DB 54 LVQLDLSF 61

RESULT 41
054910
ID 054910 PRELIMINARY: PRT: 356 AA.
AC 054910:
DT 01-NOV-1996 (TREMblrel, 01, Created)
DT 01-NOV-1996 (TREMblrel, 01, Last sequence update)
DE 01-JUN-2002 (TREMblrel, 21, Last annotation update)
DE 0119peptidase.
GN OPPD.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN 11
RC SEQUENCE FROM N.A.
RC STRAIN=SEKOTYPE M49;
RX MEDLINE=87039725; PubMed=8885277;
KA Podbielski A., Pohl B., Moisschnik M., Koerner C., Schmidt K.H.,
KA Rodzinski E., Leonard R.A.B.;
R1 "Molecular characterization of a group A streptococcal (GAS)
R1 oligopeptidase (Opp) and its effect on cysteine protease
R1 production.";
R1 Mol. Microbiol. 21:1087-1099(1996)
DR EMBL: X89237; CAA61525.1; -.
DR InterPro: IPR003593; AAA_Atpase
DR InterPro: IPR003439; ABC_Transport;
DR Pfam: PF00005; ABC_tran; 1.
DR PRODOM: PD000006; ABC_Transport; 1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; UNKNOWN_1.
KW ATP-binding; 356 AA; 39645 MW; E5670779405256A CPG64;
SQ SEQUENCE 356 AA: 39645 MW; E5670779405256A CPG64;

Query Match 2.5%; Score 8; DB 2; Length 356;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 100.0%; Pred. No. 13;

```

```

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 178 QELTDLKT 185
DB 78 QELTDLKT 85

RESULT 42
09A1F7
ID 09A1F7 PRELIMINARY: PRT: 356 AA.
AC 09A1F7:
DT 01-JUN-2001 (TREMblrel, 17, Created)
DT 01-JUN-2001 (TREMblrel, 17, Last sequence update)
DE 01-JUN-2002 (TREMblrel, 21, Last annotation update)
DE Oligopeptidase (ATP-binding protein).
GN OPPD OR SPY0296.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN 11
RC SEQUENCE FROM N.A.
RC STRAIN=SEKOTYPE M49;
RX MEDLINE=8119284; PubMed=11296296;
KA Ferrerelli J.J., M'Shar W.M., Adler T., Savir T., Savir G., Lyon K.,
KA Frenkel G., Secor S., Surooy A.N., Kention S., Lai H.S., Lin S.P.,
KA Qian Y., Liu R.G., Najjar P.Z., Fan G., Zhu H., Song L., White J.,
KA Yan X., Clifton S.W., Roe B.A., McLaughlin R.;
R1 "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
R1 Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
DR EMBL: AE006495; AAK33363.1; -.
DR InterPro: IPR003593; AAA_Atpase.
DR InterPro: IPR003439; ABC_Transport;
DR Pfam: PF00005; ABC_tran; 1.
DR PRODOM: PD000006; ABC_Transport; 1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; UNKNOWN_1.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 356 AA: 39703 MW; E565AEFP405256A CPG64;

Query Match 2.5%; Score 8; DB 10; Length 356;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 178 QELTDLKT 185
DB 78 QELTDLKT 85

RESULT 43
082594
ID 082594 PRELIMINARY: PRT: 382 AA.
AC 082594:
DT 01-NOV-1998 (TREMblrel, 08, Created)
DT 01-NOV-1998 (TREMblrel, 08, Last sequence update)
DE 01-DEC-2001 (TREMblrel, 19, Last annotation update)
DE FliG4.2 protein.
GN FliG4.2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Eudicotyledons; Core eudicots; Rosidae;
OC Eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN 11
RC SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX Washor;
R1 "The A. thaliana genome sequencing project.";
R1 Submitted (09-1998) to the EMBL/Genbank/DBJ databases.
RN 12
RC SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX Arabidopsis thaliana; Stomcking L., Larsson Y., Trevisan E.;

```

```

KT      "the sequence of A. thaliana F1104."
KL      Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=CV. COLUMBIA;
RA      Waterston R.;
RL      Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AF096470; AAC62779.1;
DR      InterPro: IPR000477; PVT5;
DR      InterPro: IPR000531; TonB_boxC;
DR      Pfam: PF00078; rvc; 1;
DR      PROSITE: PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
KW      RNA-directed DNA polymerase
SQ      SEQUENCE 382 AA: 43977 MW: 0DE2342C21DD3745 CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 10; Length 382;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      47 VPSSLSQE 54
DB      165 VPSSLSQE 172

RESULT 44
Q9C1R9          PRELIMINARY:      PRT:      441 AA.
ID      Q9C1R9          PRT:      441 AA.
AC      Q9C1R9;
DT      01-JUN-2001 (TrEMBLrel. 17, Created)
DT      01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT      01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE      Centromere binding protein 1 (Centromere binding factor 1)
GN      CBF1.
OS      Candida glabrata (Yeast) (Torulopsis glabrata)
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; mitosporic Saccharomycetales, Candida.
OX      NCBI_TaxID=5478;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      MEDLINE=21332095; PubMed=11438645;
RA      Stoyan T., Gloeckner G., Diekmann S., Carbon J.;
RT      "Multifunctional Centromere Binding Factor 1 Is Essential for
RT      Chromosome Segregation in the Human Pathogenic Yeast Candida
RT      glabrata."
RL      Mol. Cell. Biol. 21:4875-4888(2001).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Stoyan T., Gloeckner G., Diekmann S., Carbon J.;
RT      "Cloning of a centromere binding factor (CBF1) gene from Candida
RT      glabrata."
RL      Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AF233433; AAK1499.1;
DR      EMBL: AF159250; AAK58891.1;
DR      HSSP: P22415; IAN4.
DR      InterPro: IPR001092; H1H_Distric;
DR      Pfam: PF00010; H1H_1;
DR      PRINTS: PR01632; PQVDCALPHAI;
DR      SMART: SM00353; HCH; 1;
SQ      SEQUENCE 441 AA: 51409 MW: 477955CD32F9BCB0 CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 3; Length 441;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      154 YQELTRLK 161
DB      367 YQELTRLK 374

RESULT 45
Q9TVX8          PRELIMINARY:      PRT:      472 AA.
AC      Q9TVX8; 017975;

```

```

DT      01-MAY-2000 (TrEMBLrel. 13, Created)
DT      01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT      01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE      R31.3 protein (OSM-6 protein).
GN      R31.3 OR OSM-6.
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; R;
OC      Rhabditidae; Rhabditinae; Caenorhabditis.
OX      NCBI_TaxID=6239;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Harris B.R.;
RT      Submitted (MAY 1996) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RA      MEDLINE=9909613; PubMed=9851916;
RN      [3]
RP      SEQUENCE FROM N.A.
RA      STRAIN=BRISTOL N2;
RT      "Genome sequence of the nematode C. elegans: A platform
RT      for investigating biology."
RL      Science 287:2012-2018(1998).
RN      [3]
RP      SEQUENCE FROM N.A.
RA      COLLET J., SPIKE C.A., LUNDQUIST E.A., SHAW J.E., HERNAN
RT      "Analysis of osm-6, a gene that affects sensory cilium s
RT      sensory neuron function in Caenorhabditis elegans."
RL      Genetics 148:187-200(1998).
DR      EMBL: Z75954; CAB61020.2;
DR      EMBL: Z75955; CAB61020.2; JOURNAL;
DR      EMBL: Z75954; CAB61020.2; JOURNAL;
DR      EMBL: AJ000259; CAB3975.1;
SQ      SEQUENCE 472 AA: 53154 MW: 5FA6924EB1C4278F CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 5; Length 472;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      169 DQSKQVQI 176
DB      21 DQSKQVQI 28

```

Search completed: December 7, 2002, 11:16:57
Job time : 39 secs

